

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2003, 18:25:31 ; Search time 20 Seconds
(without alignments)
801.775 Million cell updates/sec

Title: US-09-926-799-1
Perfect score: 2896
Sequence: 1 MMASKDAPTMDGTSGAQ.....YQLKPVGTAGPACRLGIRRS 545

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A.COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B.COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A.COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B.COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PCTUS.COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	285	9.8	626	4	US-09-590-020-7
2	272	9.4	623	4	US-09-590-020-4
3	269	9.3	623	4	US-09-590-020-2
4	262.5	9.1	622	4	US-09-590-020-6
5	149	5.1	2206	1	US-07-852-260-2
6	149	5.1	2206	2	US-08-461-503-2
7	149	5.1	2206	4	US-08-465-250-2
8	106	3.7	1302	4	US-09-423-890-2
9	106	3.7	1493	4	US-09-423-890-8
10	106	3.7	1593	4	US-08-628-829-4
11	106	3.7	2318	4	US-09-091-219-24
12	103.5	3.6	1170	4	US-09-749-588-2
13	102.5	3.5	588	1	US-08-460-860-4
14	102	3.5	1048	4	US-09-171-699-10
15	100	3.5	544	4	US-09-615-192A-349
16	98	3.4	1209	4	US-09-749-588-4
17	97	3.3	956	4	US-09-134-078-63
18	97	3.3	3443	2	US-08-416-603-2
19	95.5	3.3	202	2	US-08-416-603-10
20	95	3.3	2232	4	US-09-091-219-25
21	95	3.3	2247	4	US-09-091-219-2
22	94.5	3.3	2004	1	US-08-375-709-15
23	94.5	3.3	2004	1	US-08-752-929-15
24	94.5	3.3	2004	4	US-09-090-793-9
25	94	3.2	1036	4	US-09-206-942-73
26	94	3.2	1477	1	US-08-038-682-4
27	94	3.2	1477	1	US-08-302-832-4

28	34	3.2	1477	2	US-08-530-198-4	Sequence 4, Appli
29	34	3.2	1477	2	US-08-469-880-4	Sequence 4, Appli
30	34	3.2	1477	2	US-08-728-470-4	Sequence 4, Appli
31	34	3.2	1477	2	US-08-617-697-4	Sequence 4, Appli
32	34	3.2	1477	4	US-08-719-641-4	Sequence 4, Appli
33	34	3.2	1477	4	US-09-206-942-71	Sequence 2, Appli
34	93.5	3.2	907	3	US-08-783-774-2	Sequence 71, Appli
35	93.5	3.2	907	4	US-09-328-599A-1	Sequence 2, Appli
36	93.5	3.2	907	5	PCT-US95-04611A-19	Sequence 1, Appli
37	93	3.2	20	4	US-08-973-961-5	Sequence 19, Appli
38	93	3.2	1381	4	US-09-540-245A-16	Sequence 5, Appli
39	92.5	3.2	669	4	US-09-071-035-264	Sequence 16, Appli
40	92.5	3.2	1638	4	US-09-071-035-258	Sequence 264, App
41	92.5	3.2	1638	4	US-09-071-035-262	Sequence 258, App
42	92.5	3.2	1638	4	US-09-071-035-266	Sequence 266, App
43	92.5	3.2	2308	1	US-08-015-973-1	Sequence 266, App
44	92.5	3.2	2308	2	US-08-448-164-1	Sequence 1, Appli
45	92.5	3.2	2308	4	US-08-081-929-2	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-09-590-020-7
; Sequence 7, Application US/09590020
; Patent No. 6355246
; GENERAL INFORMATION:
; APPLICANT: Kruger, John M
; APPLICANT: Maes, Roger K
; APPLICANT: Vilnis, Aivars
; TITLE OF INVENTION: METHOD FOR THE ISOLATION AND DIAGNOSIS OF FELINE
; TITLE OF INVENTION: CALICIVIRUS AND VACCINES THEREOF
; FILE REFERENCE: MSU 4.1-446
; CURRENT APPLICATION NUMBER: US/09/590, 020
; PRIOR FILING DATE: 2000-06-08
; PRIOR APPLICATION NUMBER: 60/138,484
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 626
; TYPE: PRI
; ORGANISM: Feline calicivirus
US-09-590-020-7

Query Match 9.8%; Score 285; DB 4; Length 626;
Best Local Similarity 23.8%; Pred. No. 4.7e-20;
Matches 144; Conservative 87; Mismatches 194; Indels; 180; Gaps 32;
QY 13 DGTSGAGQLPEANT-AEPTSMPEVACAATAAGOVNIDPMIMNMYVQAPQGEFTI- 70
DB 82 DGSITA----PEQGTWGGVIAEPAQMSAAMATKSKVDSEW-----EAFESFH 128
QY 71 ----SPNNTPGDILDLQGLPHLNPFLSLHAQMYNGVMKVKVLLAGNAFTAGRIII 125
DB 129 TSVNWSSETQGLKILFKQSLGLPLNPLYSLHLAKLYVAMSGSIEVRFSGVFGGLAA 188
QY 126 SCIDPGP-AAQNTSIAQATWPHVIAVRVLEPIEVLEDRVNVLFNNNDNAPTRMLVCM 184
DB 189 IVVPGGVDVQSTSMLO---YPHVLFDAQVPEVIFCLPDLRSLTYHLSMDTDTSTSLVIM 245
QY 185 LY----TPLRASGSSGCTDFPVIAGRYLTCPSDFSEFLVPPNVEQKTRPFSVPLNPLN 240
DB 246 VYNDLNPYANDYNSSCC---IVT--VETKPGDPFKHLLKPPG-----S 285
QY 241 TLSNRPVS-LI----KSMYMSRDH-GOMVQFNGRVRTLDGOLQGTTPPTSASOLCKIRSV 295
DB 286 MLTHGVSFSDLPKSSSLIGNRHSWDTDF-----IIRPFV 322
QY 296 PHANGNGYN-----LTELDSGYHA-----FESPA-PIGFPDL---G 329
DB 323 FQANRHFDFNOETAGWSTPRFRPISVTITEQNGAKLGIGVATDVIYVPGIDGMPDFTIPG 382

```

QY 330 EC-----DWHM-----EASPTQNTGDIK-QINVK-----QESAPA 361
Db 383 ELIPAGDAITNGTNDITATGYDTADIIKNNTNFRGMYICGSLQRAWGDKISNTAF- 441
QY 362 PHLGTIQADGLSDSVN-TNMI-----AKLGWVSPVSD-----GHRGVDVDPWI 404
Db 442 -ITATLDGNNKINFCNTIDQSKIVVFQDAHVGHKKAQTSDDTLALLGYTGIGEQAI- 498
QY 405 PRYGSTLLEAAOLA--PPIYPGGEAIVFFMSDFPIAHGTNGLSVPCCTIPOEFVTHFN 462
Db 499 ---GSDRRVVRISTLPETGARG-GNHPFIYKNSIKLGYVIRSIDV-----FNS 543
QY 463 EOAPTRGEAALLHYLDP-----DTHRNLFGEFKLYPEGEFTCVPNSSGTGPTQLPI 512
Db 544 QILHTRSLNLNHLPLPDPSFAVYRIIDNSNGSDFDIDGIDGDF--SFVGVSGFGKLEFPL 601
QY 513 NGVFV 517
Db 602 SASYM 606

RESULT 2
US-09-590-020-4
; Sequence 4, Application US/09590020
; Patent No. 6355246
; GENERAL INFORMATION:
; APPLICANT: Kruger, John M
; APPLICANT: Maes, Roger K
; APPLICANT: Vilnis, Alvars
; TITLE OF INVENTION: METHOD FOR THE ISOLATION AND DIAGNOSIS OF FELINE
; TITLE OF INVENTION: CALICIVIRUS AND VACCINES THEREOF
; FILE REFERENCE: MSU 4.1-446
; CURRENT APPLICATION NUMBER: US/09/590,020
; CURRENT FILING DATE: 2000-06-08
; PRIOR APPLICATION NUMBER: 60/138,484
; PRIOR FILING DATE: 1999-06-10
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 4
; LENGTH: 623
; TYPE: PRT
; ORGANISM: Feline calicivirus
US-09-590-020-4

Query Match 9.4%; Score 272; DB 4; Length 623;
Best Local Similarity .22.2%; Pred. No. 9.7e-19;
Matches 133; Conservative 84; Mismatches 209; Indels 174; Gaps 27;

QY 7 DAPT-NMDGTSGAGOLVPEANTAEPTISMEPVAGAATAAATAGQVNMIDPWIMNNYVQAPQ 65
Db 72 DLPLERLEGDDGSIITTEQGTWVGVIAPSAQMSAAADMATGKSVDSW-----E 122
QY 66 GEFTI-----SPNNTPGDILFDLQGLHLPFLSHLAQMYNGWGNMKVKVLLAGNAFT 119
Db 123 AFFSFTSVNWSSTQGTWVGVIAPSAQMSAAADMATGKSVDSW-----E 122
QY 120 AKIIISCIPPGF-AAONISIAQATMFPVHIVADRVLEPIEVPLEDVRNVLHNNDAFT 178
Db 123 AFFSFTSVNWSSTQGTWVGVIAPSAQMSAAADMATGKSVDSW-----E 122
QY 120 AKIIISCIPPGF-AAONISIAQATMFPVHIVADRVLEPIEVPLEDVRNVLHNNDAFT 178
Db 183 GGLAAIVVPPGIEPVQSTSMQ---YPHVLFDAQVPEVIFAIPDLRSNLYHLMSDSTD 239
QY 179 MRLVCMLY-----TPLRAGSSSGTDFPVIAGRVLTCPSPDFSLFLVPPNVEQKTPFSV 234
Db 240 TSLVIMVYNDLIPYANDTNSGSC---IVT--VETKPGDPKFKHLLKPPG----- 284
QY 235 PNLPLNTLSNRVPS-LI-----KSMVYSRQHGOMVQ-----FQ-NGRVTLDGQLQG-- 279
Db 285 -----SMLTHGVSVDLPKSSSLTGNRHSWSDITFIIRPFVQANRHDFDNOETAGWS 339
QY 280 -----TPTPSAQLCKI-----RGSVFHANGNGNYN 305
Db 340 TPRFRPITITVSESNKSLGIGVATDVIYVPGWPDWTIPEQLTPAGIYSITASNGTD 399

```

```

QY 306 LTLEDGSPYHAFSPA-PIGFPDLGEC-----WHMEASPTTQF-----NT 345
Db 400 ITTAAG--YDAAETIVNTNFKSMYICGSLQRAWGDKISNTAFITTAVRKGNISIEPSNT 457
QY 346 GDVIKQINVKQESAFAPHLGTIQADGLSDSVSVNTNMIKLGWVSPVSDGHRGVDVP- 401
Db 458 IDMTKLIV-VYODA-----HVG-----EEVQTSDIITLALLGYTGIGEEAIGSDRQKVRI 505
QY 402 WVLPRYGSTLLEAAQLAPPIYPGGEAIVFFMSDFPIAHGTNGLSVPCCTIPOEFVTHFN 461
Db 506 SVLPETGAR-----GNHPIFYKNSIKLGYVIRSIDV-----FN 539
QY 462 NEOAPTRGEAALLHYLDP-----DTHRNLFGEFKLYPEGEFTCVPNSSGTGPTQLPI 511
Db 540 SQILHTRSLNLNHLPLPDPSFAVYRIIDNSNGSDFDIDGIDGDF-----SFVGVSNLP 592

RESULT 3
US-09-590-020-2
; Sequence 2, Application US/09590020
; Patent No. 6355246
; GENERAL INFORMATION:
; APPLICANT: Kruger, John M
; APPLICANT: Maes, Roger K
; APPLICANT: Vilnis, Alvars
; TITLE OF INVENTION: METHOD FOR THE ISOLATION AND DIAGNOSIS OF FELINE
; TITLE OF INVENTION: CALICIVIRUS AND VACCINES THEREOF
; FILE REFERENCE: MSU 4.1-446
; CURRENT APPLICATION NUMBER: US/09/590,020
; CURRENT FILING DATE: 2000-06-08
; PRIOR APPLICATION NUMBER: 60/138,484
; PRIOR FILING DATE: 1999-06-10
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2
; LENGTH: 623
; TYPE: PRT
; ORGANISM: Feline calicivirus
US-09-590-020-2

Query Match 9.3%; Score 269; DB 4; Length 623;
Best Local Similarity 29.8%; Pred. No. 2e-18;
Matches 76; Conservative 41; Mismatches 94; Indels 44; Gaps 9;

QY 7 DAPT-NMDGTSGAGOLVPEANTAEPTISMEPVAGAATAAATAGQVNMIDPWIMNNYVQAPQ 65
Db 72 DLPLERLEGDDGSIITTEQGTWVGVIAPSAQMSAAADMATGKSVDSW-----E 122
QY 66 GEFTI-----SPNNTPGDILFDLQGLHLPFLSHLAQMYNGWGNMKVKVLLAGNAFT 119
Db 123 AFFSFTSVNWSSTQGTWVGVIAPSAQMSAAADMATGKSVDSW-----E 122
QY 120 AKIIISCIPPGF-AAONISIAQATMFPVHIVADRVLEPIEVPLEDVRNVLHNNDAFT 178
Db 183 GGLAAIVVPPGIEPVQSTSMQ---YPHVLFDAQVPEVIFAIPDLRSNLYHLMSDSTD 239
QY 179 MRLVCMLY-----TPLRAGSSSGTDFPVIAGRVLTCPSPDFSLFLVPPNVEQKTPFSV 234
Db 240 TSLVIMVYNDLIPYANDTNSGSC---IVT--VETKPGDPKFKHLLKPPG----- 284
QY 235 PNLPLNTLSNRVPS 249
Db 285 -----SMLTHGVSVP 294

RESULT 4
US-09-590-020-6
; Sequence 6, Application US/09590020
; Patent No. 6355246
; GENERAL INFORMATION:
; APPLICANT: Kruger, John M
; APPLICANT: Maes, Roger K
; APPLICANT: Vilnis, Alvars

```

; TITLE OF INVENTION: METHOD FOR THE ISOLATION AND DIAGNOSIS OF FELINE
; FILE OF INVENTION: CALICIVIRUS AND VACCINES THEREOF
; CURRENT APPLICATION NUMBER: US/09/590,020
; PRIOR FILING DATE: 2000-06-08
; PRIOR APPLICATION NUMBER: 60/138,484
; PRIOR FILING DATE: 1999-06-10
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 622
; TYPE: PRT
; ORGANISM: Feline calicivirus
US-09-590-020-6

Query Match 9.1%; Score 262.5; DB 4; Length 622;
Best Local Similarity 22.3%; Pred. No. 8.9e-18;
Matches 134; Conservative 82; Mismatches 213; Indels 171; Gaps 28;
QY 7 DAPT-NMDGTSGAGQLVPEANTAEPISEMPVAGATAAATAGQVNMIDPHIMNYYVOAPQ 65
DB 72 DLPLEREGDDGSIITPEQGTWGGVIAEPAQMSAADMATGKSDVSEW-----E 122
QY 66 GEFTI-----SPNNTPGDILFDLQGLPHLNPFLSHLAQMYNGWGNKVKVLLAGNAFT 119
DB 123 AFFSFHTSVNMWSTSETQGKILFKQSLGPLLNPYLSHLAKLYVALAGSVREFSGSGVF 182
QY 120 AGKIIISCIPPGF-AAONISIAQATMPHVIADRVRLVLEIEVPLEDVNRVLFHNNDAPT 178
DB 183 GGLAAIIVPPGIEPVQSTSMQ---YPHFLDARQVPIFAIPDRLSNLYHLMSDITD 239
QY 179 MRLVCMLY----TPLRASGSSGTDPEVIAGRVLTCPSPDFSELFVPPNVEQKTFPSV 234
DB 240 TSLVIMVYNDLINPYANDTNSGC---IVT--VETKPGDPKFHLKPPG----- 284
QY 235 PNLPLNTLSNRVPS-LI---KSMVSRDHQOMVQ-----FQ-NGRVTLDCQLQ-- 279
DB 235 ----SMLTHGSPVSDLPKSSSLWGNRHWSIDITFIIRPFVFOANRHFDFNOETAGWS 339
QY 230 -----TTPTASALCKI-----RGSVFHANGNGYN 305
DB 340 TPRFRPTITVSESNMKSGLIGVATDVIYVPIPDGWPDTTIPQLTPAGIYSITASNGTD 399
QY 306 LTELDDGSPYHAFESPA-PIGFPDLGEC-----WHMEASPTTQF-----NT 345
DB 400 ITTAAG--YDAETIVTTFKSMYICGSLQRAWDKKSINTAFITTAVRKGSNIEPSNT 457
QY 346 GDVTKQINVKQESAFAPHLCTIQADGLSDVSVNTNMIKLGWYSPVSDGHRGDVDPWV-I 404
DB 458 IDMTKLV-VYQDA---HVG-----EEVQTSIDTLALLGYTGIGEEAIGSDRDKVRI 505
QY 405 PRYGSTLTEAAQLAPPYIPPGEGEAIYFFMSDFPIAHGTNGLSVPCPTIPOEFVTHFVNEQ 464
DB 506 SVLGETGARG-----GNHPIFYKNSIKLGVIRSIDV-----FNSQI 542
QY 465 APTRGEAALHY-LDPDTHR-----NLGEFKLYPEGFMTCVPNNSGTGPTLP 511
DB 543 LHSSRQLSLNLYLLPPDSFAVYRLDLSNGSWFDIGIDTEPDPG-----SFGVSNLP 595

RESULT 5
US-07-852-260-2
; Sequence 2, Application US/07852260
; Patent No. 5525715
; GENERAL INFORMATION:
; APPLICANT: Racaniello, Vincent
; APPLICANT: Tatem, Joanne M.
; APPLICANT: Weeks-Levy, Carolyn L.
; TITLE OF INVENTION: METHODS FOR PRODUCING RNA VIRUSES FROM
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham

; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10112
; COMPUTER READABLE FORM: disk
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/852,260
; FILING DATE: 19920619
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 36607-B-PCT-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 977-9550
; TELEFAX: (212) 664-0525
; TELEX: 422523 COOP UI
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2206 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-852-260-2

Query Match 5.1%; Score 149; DB 1; Length 2206;
Best Local Similarity 20.3%; Pred. No. 2.5e-05;
Matches 97; Conservative 57; Mismatches 196; Indels 128; Gaps 19;

QY 15 TSCAGOLVPEANTAEPISEMPVAGATAAATAGQV-NM-----IDPWIMNYYVQAPQG-- 66
DB 347 TFCGNYLTSNDHQSCAI--PEFDVTPPIDPGEVNRMMELAEIDTWIPLNLESTKRNTM 405
QY 67 ---EFTISPNNTPGDILFDLQGLPHLNPFLSH-----LAQMYNGWGNKVKVLLAGNAF 118
DB 406 DMVRYTLSDSADLSQPTILCLSLSPAFDRLSHTMLGEVLNYYTHWAGSLKFTFLFCGSM 465
QY 119 TAGKIIISCIPPGFAAQNISIAQATMPHVIADRVRLVLEIEVPLEDVNRVLPVH--NNDNA 176
DB 466 ATGKIIIVAPPG-AQPPTSRKEAMLGTHVINDLGLQSSCTMVVPWISNVTYRQTQDSF 524
QY 177 PMRLVCMLY----TPLRASGSSGTDPEVIAGRVLTCPSPDFSELF-------V 221
DB 525 TEGGYISMFYQTRIVVYPLSTPKSMS-----MLGFVSAC--NDFSRLLRDTHISQSAL 576
QY 222 PNVEQKTK-----PFSVPNLP-----LNTLSNRVPSLIKSM 255
DB 577 PCGIEDLTSEVAGCALTLSPKQDLSLPTDKASGPAHSKEVPALTAVETGATNPLAPSDT 636
QY 256 VSRDHQOMVQ-----FQNGRVTLDCQLQCTTPTSASQLCKIRGSVFH----- 297
DB 637 VCTRH--VVQRRSRSESTIESFARGACVAIIEVDNEQPTTTRAKQLFAMWIRITYKDTVOL 694
QY 298 -----ANGNGYNLTEDGSPYHAFESPAPIGFPDLGECOW 333
DB 695 RRKLEFFTYSRFDMFTFVVVTANFTNANNGHALNQVYQIMYIPPGAPTPKSDDY---TW 751
QY 334 HMEASPTTQFNTGDVTKQINVKQESAFAPHLGTIOA-----DGLSDVSVNTNMIKLG 386
DB 752 QTSNPSNFIYTYGAAPARISV-----PYVGLANAYSHFYDGFQAKVPLKTDANDQIG 802

RESULT 6
US-08-461-503-2
; Sequence 2, Application US/08461503
; Patent No. 5834302
; GENERAL INFORMATION:
; APPLICANT: Racaniello, Vincent

Db 943 RTIIVP-PLK-TOLGDCATVATQASGLLSNKTKPVASVSGSSCCITPTGTGQRAQRGGTSA 1000
QY 374 ----DVSNTMTAKLGWSPVSDGHRGDVDPWIPRYGTLTEAAQLAPPYPPGFGGEA 429
Db 100- AOPNLNLSONQSSA-----APTQERSNPAP-----RQQAFAVAP-----LSQA 1040
QY 430 IVFMSDFPIAHGTNG---LSVPCTIPQEFVTHFVNEQAPTGCEA-----ALLHYLDPD- 480
Db 1041 PYTFQHSPL-HSTGPHLAPAPAHLPQ--AHLYTYAAPTSAALGSTSSIAHLESPOG 1097
QY 481 THRLNGEFKLYPEGFMTCPVNSSGTGPQTL 510
Db 1098 SSRHAAAYTHPTSLVHQVPVS--VGPSLL 1125
RESULT 13
US-08-460-860-4
; Sequence 4, Application US/08460860
; Patent No. 5665584
; GENERAL INFORMATION:
; APPLICANT: HATAMOTO, OSAMU
; APPLICANT: WATARAI, TERUO
; APPLICANT: MIZUSAWA, KIYOSHI
; APPLICANT: NAKANO, EIICHI
; TITLE OF INVENTION: A DNA FRAGMENT CONTAINING A TANNASE
; TITLE OF INVENTION: GENE, A RECOMBINANT PLASMID, A PROCESS FOR PRODUCING
; TITLE OF INVENTION: TANNASE, AND A PROMOTER
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, FOURTH FLOOR
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/460,860
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 159973/1994
; FILING DATE: 12-JUL-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 83973/1995
; FILING DATE: 10-APR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F
; REFERENCE/DOCKET NUMBER: 7127-001-0X
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 413-3000
; TELEFAX: (703) 413-2220
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 588 amino acids
; TYPE: amino acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-460-860-4
Query Match 3.5%; Score 102.5; DB 1; Length 588;
Best Local Similarity 21.7%; Pred. No. 0.15;
Matches 66; Conservative 41; Mismatches 100; Indels 97; Gaps 18;
QY 273 LDGQLQGTPTTSASQLKIRGSVFHANGNGYNLTDLGDSYH-AFESPAPGPDIGEC 331
Db 264 LDGRDGV--VSRDLCKL-----NFLTITIGEPYCAAGTSTSLGF----- 304

QY 332 DHMEASPTTQNTGDVLIKQINVKOE-----SAPAP-HLGTQADGLSDVSV-----N 378
Db 305 -----GFSNG---KRSNVKQAECSITSYQPAQNGTGTARGVAVAQAIYDGLHNS 351
QY 379 TMMIAKLGW--VSPVSDG---HRGDVDPWV---IPRYGS-----TLTEAAQLAPPY 421
Db 352 KGERAYLSHQIASELSDAETENSDTGKWEINIPSTGGGYVTKFIQLLNLDNLNNVT 411
QY 422 YPPGGEAIVFFMSDFPIAHGTNGLSVPCTIPQ-----EFVTHFVNEQAPTGCEAAL 473
Db 412 Y-----DTLVDWMTGMVRYMD---SLQTTLPDLTPFOSSGGKLLHYHGESDPSIPAASS 463
QY 474 LHYLDP-----DTHRLNGEFKLYPEGFMTCPVNSSGTGPQTLPIGVFVF 518
Db 464 VHYWQAVRSVMYGDKTEEEALEALEDWYQFYLLP-GAAHCGTNSLQPGP--YPENNMMEIM 520
QY 519 VSKV 522
Db 521 IDWV 524
RESULT 14
US-09-171-699-10
; Sequence 10, Application US/09171699
; Patent No. 6448389
; GENERAL INFORMATION:
; APPLICANT: The Wistar Institute of, Anatomy & Biology
; Gonczol, Eva
; Berencsi, Klara
; Kari, Csaba
; TITLE OF INVENTION: No. 6448389el Cytomegalovirus DNA Constructs and
; Uses Therefor
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howson and Howson
; STREET: Spring House Corporate Center, P.O. Box 457
; CITY: Spring House
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19477
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/171,699
; FILING DATE: 19-Jan-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/015,717
; FILING DATE: 23-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Kodroff, Cathy A.
; REGISTRATION NUMBER: 33,980
; REFERENCE/DOCKET NUMBER: WST66APCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-540-9200
; TELEFAX: 215-540-5818
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1048 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 10:
US-09-171-699-10
Query Match 3.5%; Score 102; DB 4; Length 1048;
Best Local Similarity 21.8%; Pred. No. 0.44;
Matches 120; Conservative 50; Mismatches 214; Indels 166; Gaps 25;

QY 15 TSGAGQLVPEANTAPRISMEPV-----AGAATAAATAGQVNMIDPWNMNNVVOAPQG 66
 Db 578 TEGAGVNVNVPAGAAAILTPTVNESTAPAPAPTPTTFACTQTPVNGNSPW-----APTA 630
 QY 67 EF--TISPNTTGDILFQLQGLPHL--NPFSLHQAQVNGVGNMKVKVLGAGNAFTAGK 122
 Db 631 PLPGDMNANPREAWALK-NPHLAYNPFMPPTS-----TASQ 669
 QY 123 IISCIPGFAAQNTISIAQATMFHVIADRVL--EFIEVPLEDVNRVLFHNNDAFTMR 180
 Db 670 NVTSTPRPSTPRAAVTQ-TASRDAADDEVWALRDQTAESVEDSE-----714
 QY 181 LVCMLYTLRASGSSGSDPFIAGRVLTCPSPDFSEFLVPPNVEOKTKPESVP-----235
 Db 715 -----EEDDDSDTGSVVSLGH--TTPSSDYNNDVISPPS---QTPEQSTPSIRK 760
 QY 236 ----NPLNTLSNSRVPISLIKSMVSRDH---GOMVQFQNGRVTLTGQLOGTTPTSASQL 288
 Db 761 AKLSSPMTTTSTQKPVLGK--RVATPHASARAQTVTSTPVQGRLEKQVSGTPTVPATL 818
 QY 289 CKIRGSVHANGNGNYLTGDSYPHAFESPAPIGPFDLGCEDHMEASPTTQFNT---345
 Db 819 LQ-----PQP-----ASSKTTSSRNVT 836
 QY 346 -GDVVKOINVKESAFAPHLGTIOAGLSVSVNTNMI-----AKLGNVSPVSD-GHR 396
 Db 837 GAGTSSASARQPSASASVLSPTEDDVVSPATSPLSMLSSASPSAKSPAPSPVKGRGRS 896
 QY 397 GDVDWVPIRGYSTLEAAQLAPPIYP-----PFGGAIVFFMSDFPIAGTNCLSVPCT 451
 Db 897 VGV-PSLKRPTLGG---KAVGRRPSPVPSVGSAPGRSLSSRAASTTPTPAVTTVYPPSS 952
 QY 452 IQPEFVTHFVNQAAPT---RGEAALLHYLDPTDTHRNLEKFEK-----LYPEGFMT 498
 Db 953 TAKSSVSNAPPVAPSPILKPGASAL-----QSRRTCTRAVGSVPKSTTGKMTKVAFDLS 1007
 QY 499 VNSSGTGPQ 508
 Db 1008 SPQKSGTGPQ 1017

RESULT 15
 US-09-615-192A-349
 ; Sequence 349, Application US/09615192A
 ; Patent No. 6410718
 ; GENERAL INFORMATION:
 ; APPLICANT: Bloksberg, Leonard N.
 ; APPLICANT: Havukkala, Ilkka
 ; TITLE OF INVENTION: Materials and Methods for the
 ; FILE REFERENCE: 11000.100364U
 ; CURRENT APPLICATION NUMBER: US/09/615.192A
 ; CURRENT FILING DATE: 2000-07-12
 ; PRIOR APPLICATION NUMBER: US 08/975,316
 ; PRIOR FILING DATE: 1997-11-21
 ; PRIOR APPLICATION NUMBER: US 08/713,000
 ; PRIOR FILING DATE: 1996-09-11
 ; PRIOR APPLICATION NUMBER: US 09/169,789
 ; PRIOR FILING DATE: 1998-10-09
 ; NUMBER OF SEQ ID NOS: 405
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 349
 ; LENGTH: 544
 ; TYPE: PRT
 ; ORGANISM: Pinus radiata
 US-09-615-192A-349

Query Match 3.5%; Score 100; DB 4; Length 544;
 Best Local Similarity 19.9%; Pred. No. 0.24;
 Matches 109; Conservative 68; Mismatches 197; Indels 174; Gaps 30;
 QY 7 DAPTNMDGTSGAGQLVPEANTAPRISMEPVAGAAATAAATAGQVNMIDPWNMNNVVOAPQG 66

Db 42 DRPCVINGATORITYYAE---VELISRRVSAGNLGVLGGQGVIML---LLQN---CPEF 92
 QY 67 EF-----TISPNTTGDILFQLQGLPHLNFPIHSLAQMYN 101
 Db 93 VFAFLGASVYRGAISTANPEVTPGCEIAKQASAAKIVITQAAAFADKVRFAEE---N 147
 QY 102 GWVGNMKVKVLLAG-----NAFTAGKI---IISCIPGFAAQNTISIAQ 141
 Db 148 G-----VKVVCIDTAPEGCLHFSLMQADENAAAPADVPDDVLALP--YSSGTTGLPK 199
 QY 142 ATMFPHVIADRVLEPIEVPLEDVNRVLFHNNND---NAPTMRLVCM---LYTPLRASGS 194
 Db 200 GYMLTHRGQVTSVAQVD---GDNPNLYFFHKEDVILCTPLPLFHIYSLSNVMFCALRVGAA 256
 QY 195 SSGTDFVIAGRVLTCPSPDFSEFLVPPNVEOKTKPFSVPLNPLNTLSNSRVPISLIKSM 254
 Db 257 ILMOKFEIVALMELVQRYRYVTILPIVPPVILEIAKSAEVDYDL-----SSIRTI 307
 QY 255 MVSRDHGQMVQFQNGRVTLTGQLOGTTPTSASQLCKIRGSVHANGNGNYLT-----LD 310
 Db 308 M-----SCAAPMGKELEDT-----VRAKLPNAKLGQGYCMTGAGPVLA 345
 QY 311 GSPYHAFESPAPIGPFDLGCEDHMEASPTTQFNTGVIVKQINVK-----QESAPAP--HL 364
 Db 346 MCPAPAKE-PFEI---KSGAC-----GTVVRNAEMKIYDPETGASLPRNOA 387
 QY 365 GTIQADG-----LSDVSVNTNMTAKLGWVSPVSDGHRGDV-----DPWVPIRGYST 410
 Db 388 GEICIRGHQIMKGYLNDAAEATANTIDKEGWL-----HTGDIGYIDDDDELFIYDRKEL 441
 QY 411 LT-EAAQLAPP-----IYPPGGAIVFFMSD-----FPIAH--GTNGLSVPTCTPQEF 456
 Db 442 IKYKGFQVAPARLEAMLIHAHPSISDAAVPMKDEVAGEVPVAVVVKSN-----SVITEDE 497
 QY 457 VTHFVNEQ 464
 Db 498 IKQYISKQ 505

Search completed: March 10, 2003, 19:00:44
 Job time : 26 secs

GenCore version 5.1.3
Copyright (C) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2003, 18:25:00 ; Search time 23 seconds
(without alignments)
2277.968 Million cell updates/sec

Title: US-09-926-799-1

Perfect score: 2896

Sequence: 1 MMASKDAPTNMDGTSAGQ.....YQLKPVGTAGPACRLGIRRS 545

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_73:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	1974.5	68.2	546	2 B37491	major capsid prote
2	1946	67.2	530	2 B37471	capsid protein - N
3	1223	42.2	542	2 S60616	capsid protein - h
4	1151.5	39.8	539	2 S40111	capsid protein - h
5	316	10.9	2344	1 RWRH	genome polyprotein
6	315	10.9	576	2 A53982	capsid protein - E
7	311.5	10.8	2344	2 S53399	genome polyprotein
8	307.5	10.6	2344	2 S64740	genome polyprotein
9	300.5	10.4	702	1 A48562	coat protein - San
10	277.5	9.6	668	2 JQ2354	capsid protein - f
11	276	9.5	671	1 VCMWF9	coat protein - fel
12	274	9.5	668	1 VCMWFF	coat protein - fel
13	269	9.3	668	1 VCMWFF	coat protein - fel
14	264	9.1	703	1 C48562	coat protein - San
15	261.5	9.0	668	2 JQ2356	capsid protein - f
16	147	5.1	2206	2 S03822	genome polyprotein
17	146.5	5.1	2205	1 GNNY2W	genome polyprotein
18	145.5	5.0	2207	1 GNNY5P	genome polyprotein
19	145	5.0	2206	1 GNNY4P	genome polyprotein
20	143.5	5.0	2194	1 GNNY7	genome polyprotein
21	141.5	4.9	2207	2 S09553	genome polyprotein
22	139.5	4.8	2206	1 GNNY27	genome polyprotein
23	137.5	4.7	2209	1 GNNY3P	genome polyprotein
24	136.5	4.7	2209	1 GNNY2P	genome polyprotein
25	135.5	4.7	2207	1 GNNY1P	genome polyprotein
26	133.5	4.6	2179	1 GNNYHA	genome polyprotein
27	128.5	4.4	613	2 T35828	acetolactate synth
28	125	4.3	3473	1 A46112	genome polyprotein
29	125	4.3	3473	2 S27927	polyprotein - rice

30 124 4.3 2214 1 A48548 genome polyprotein
31 123.5 4.3 2175 1 GNNYBE genome polyprotein
32 122.5 4.2 3085 2 T00327 polyprotein - infe
33 122 4.2 3263 2 E82410 hypothetical prote
34 117.5 4.1 2185 1 JQ2021 genome polyprotein
35 117 4.0 2183 1 GNNYB4 genome polyprotein
36 116.5 4.0 2206 1 GNNY21 genome polyprotein
37 116 4.0 1037 2 T13350 genome polyprotein
38 115 4.0 2185 1 GNNYSV genome polyprotein
39 115 4.0 2185 1 GNNYSH genome polyprotein
40 113.5 3.9 2164 1 GNNY89 genome polyprotein
41 113 3.9 833 2 S45041 genome polyprotein
42 111.5 3.9 516 2 S28060 serum response fac
43 111 3.8 3624 2 AD0835 large repetitive p
44 110.5 3.8 5188 2 B85547 probable RTX famil
45 110.5 3.8 5291 2 F90696 hypothetical prote

ALIGNMENTS

RESULT 1

B37491

major capsid protein [similarity] - Southampton virus

N:Alternate names: orf2 protein

C:Species: Southampton virus

C>Date: 22-Sep-1993 #sequence_revision 25-Apr-1997 #text_change 18-Aug-2000

C:Accession: B37491

R:Lambden, P.R.; Caul, E.O.; Ashley, C.R.; Clarke, I.N.

Science 259, 516-519, 1993

A:Title: Sequence and genome organization of a human small round-structured (Norwalk-

A:Reference number: A37491; MUID:93142023; PMID:8380940

A:Accession: B37491

A:Status: not compared with conceptual translation

A:Molecule type: nucleic acid

A:Residues: 1-546 <LAW>

A:Cross-references: GR:L07418; NID:q1236787; PIDN:AAA92984.1; PID:q295114

A:Note: sequence extracted from NCBI backbone (NCBI:123458)

A:Note: small round-structured virus, SRSV, Norwalk virus, Norwalk-like virus, seroty

C:Superfamily: human calicivirus capsid protein

C:Keywords: glycoprotein

F:303,340,441/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match

Best Local Similarity 68.2%; Score 1974.5; DB 2; Length 546;

Matches 357; Conservative 75; Mismatches 96; Indels 13; Gaps 7;

Qy 1 MMASKDAPTNMDGTSAGQLVPEANTAEPIISMPEVAGAAATAAGQVMIDPWIMNNY 60

Db 1 MMASKDAPQSDAGSAGAGQLVPEVNTADPLPMEPVAGPTTAVATAGQVMIDPWVNF 60

Qy 61 VQAPQGEFTISPNTPGDIQLGPHLNPFLSHLAQMYNGWGNMVKVLLAGNAPTA 120

Db 61 VQSPQGEFTISPNTPGDIQLGPHLNPFLSHLSQMYNGWGNMVRILLAGNAPSA 120

Qy 121 GKILISCLIPGFAONISIAQTAEPIISMPEVAGAAATAAGQVMIDPWIMNNY 180

Db 121 GKILISCLIPGFAONISIAQTAEPIISMPEVAGAAATAAGQVMIDPWIMNNY 180

Qy 121 GKILISCLIPGFAONISIAQTAEPIISMPEVAGAAATAAGQVMIDPWIMNNY 180

Db 121 GKILISCLIPGFAONISIAQTAEPIISMPEVAGAAATAAGQVMIDPWIMNNY 180

Qy 181 LVCMLYTPLRASGSSGTDPFVIAGRVLTCPSPDFSLFLVPPNVEQTKPFSPNLPUN 240

Db 181 LVCMLYTPLRASGSSGTDPFVIAGRVLTCPSPDFSLFLVPPNVEQTKPFSPNLPUN 240

Qy 241 TLSNSRVPSLIKSMVSRDHQMYQFNGRVTLDQGLGTTPTSASQLCKIRGSVFHANG 300

Db 241 TLSNSRVPSLIKSMVSRDHQMYQFNGRVTLDQGLGTTPTSASQLCKIRGSVFHANG 300

Qy 301 GNGYNLTGDSPPYHAFESPAPIGFPPDLGECDDHWEASPT-TQNTGDVIGQINVKOB-S 358

Db 301 GNGYNLTGDSPPYHAFESPAPIGFPPDLGECDDHWEASPT-TQNTGDVIGQINVKOB-S 358

Qy 299 ARTUNLTETVDGKPFMAFDPVGFPPDFGKCDWHMIRISKTPTNNTGSGDPMRSVSQTVNQ 358

Db 299 ARTUNLTETVDGKPFMAFDPVGFPPDFGKCDWHMIRISKTPTNNTGSGDPMRSVSQTVNQ 358

Qy 359 AFAPHLGTTIQADGLSDVSVNTNMIATKLGWSPVSDGHRGVDVDPVPIPRYSGTFLTEAAQLA 418

Db 359 AFAPHLGTTIQADGLSDVSVNTNMIATKLGWSPVSDGHRGVDVDPVPIPRYSGTFLTEAAQLA 418

Db 1826 QVDDQETWRTNFY---NDVFTWSVADAPGSGILYTVQHSPPNNPFTAVLSQMYAGWAGGM 1882

Qy 108 KVKVLLAGNATAGKIIITSCIPPGFAAQNISIA---QATMFPHVIAADVRLVLEPIEVPLED 164

Db 1883 QFRIVAGSGVFGGLVRAVPPG-----IEIGPGLEVRFQFPHVVIDAKSLSEPVTTMPD 1937

Qy 165 VRNVLFH--NNDNAPTRMLRVLCMLYTPL--RASGSSSGCTDPFVIAGRVLTCPSPDFSEFLV 221

Db 1938 LRPNMYHTKTPGPGLVPTLVLSYNNLINFPGSGTS-----AIQVTVETRPSDEFEFVMIR 1992

Qy 222 PPAVEQKTKPFSPVNPPLPLNTLSNRVPSLIKSMWYSRDRHGQMVQFO-----NGRVT 272

Db 1993 AP3--SKTVDSISPAGLLTT-----PVLTGVGNDRNRWGQIVGLQVPVGGFSTCRNHNW 2044

Qy 273 LDGGLQGTTPTSASOLCKIRGSHVANGNGYNLTTEL-----DGSPTYHAFESP----APIG 324

Db 2045 LNSTYGVWSSPREGDIDHRRGSASY--SGSNATNLVQFWYANAGS---AIDNPISQVAPDG 2100

Qy 325 FPDGLGCDWHHEASPTTQNTGDTVTKQINVKOESAPAPHLGTLIQADGLSDSVSYNTNNIAK 384

Db 2101 FPDMSFEPFNGPGIPAGWVGFGAIWNSN-----SGAPNWTTVQA-----YE 2142

Qy 385 LGWVSPVSDGHRGDVDPWVPIPRYGSTLTLEAAO--LAPPIYPPGGEA 429

Db 2143 LGFAT-----GAPGNLQP-----TTNTSGAOTVAKSIYAVVTGTA 2177

RESULT 6

A53982

C: capsid protein - European brown hare syndrome virus

C: Species: European brown hare syndrome virus

C: Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 28-Jul-2000

C: Accession: A53982

R: Wirblich, C.; Meyers, G.; Ohlinger, V.F.; Capucci, L.; Eskens, U.; Haas, B.; Thiel

J. Virol. 68, 5164-5173, 1994

A: Title: European brown hare syndrome virus: relationship to rabbit hemorrhagic disease virus

A: Reference number: A53982; MUID: 94309183; PMID: 7518531

A: Accession: A53982

A: Status: preliminary

A: Molecule type: mRNA

A: Residues: 1-576 <WIR>

A: Cross-references: GB: U09199

C: Superfamily: human calicivirus capsid protein

	Query Match	Best Local Similarity	20.9%;	Score 315;	DB 2;	Length 576;
	Matches 126;	Conservative	69;	Mismatches 200;	Indels	86;
	Gaps	20;				
Qy	7	DAP-----TNDMGTSGAGOLVPEANTAEPISEMPVGAATAATAGQVNMDP--WIMN	58			
				:	:	:
				:	:	:
Db	8	DAPGTTATASVPGITTDGMDPCVASTDVVTADNVAASVATAGIGGPPQQAQESWRVN	67			
				:	:	:
				:	:	:
Qy	59	NYVOAPQGEFTISPNTPGDILFDLQGLPHLPFI ⁺ LSHLAQM ⁺ YNGWYGNMKVKVLLAGNAF	118			
				:	:	:
				:	:	:
Db	68	FFY---NDVETWSVTDAPGSI ⁺ LYSVQHSQNNPFI ⁺ TOVL ⁺ SQM ⁺ YAGWAGGMGFRIVAGSGI	124			
				:	:	:
				:	:	:
Qy	119	TAGKTIISCTIDGFAAO--NISTAQATMEPHVTAADVEPI ⁺ ENVRNVLPI ⁺ -NDNVA	176			
				:	:	:
				:	:	:

```

125 FGRLVCAIIPPGIQLQGLEVRQ---FHVVIDARSLPEVITITPDLPREMIHPTGDPG 181
177 PTMLRLVCMLYTPLRASGSSGTDPF-----VIAGRVLTCPSDFSLFLVPPNVEQTKRP 231
182 LVPTLVVSVNNL-----INPFGTTSIAIQVTVETRPSEDFEVLIRAFPS--SKTVD 231
232 FSVPNPLNTLSNRVPSLIKSMVSRDHQGMVOFQ-----NGRVTLOGQLGTTTP 282
233 SVNPSMLLT-----PVLTGAGSDNRWAGAPIVGLQVPVPGFSTSNRNMNMGSTYGWSS 285
283 TSASQLCKIRGSVFHANGNGYNLTDELSPYHAFESP-----APIGFPDLGECDDHMEAS 338
286 PREDDIDHSGSNVSYPTGSAINTIETWYANAGTATTNPISNAPDGFPPDMGAIPF-----S 341
339 PTTQFNTGDIKQINVKQESAFAPHLGTIQADGLSDSVSVNTNMIKLGWVSPVSDGHRGD 398
342 GTT-ITPGAWGVGQVNNASNGTTPYVGTQA-----YELGF---ANGAPSS 383
399 VDPWVPIRYGSTLTEAAQL-APPYPPGGE-----AIVFMSDFPIAHGTNGLSVPCCTIP 453
384 IRP-----VTTTGAQLVAKSIYGVAIAQNOSSAGIIFLSKGMVS--TPGVAATTYP 434
454 Q 454
435 Q 435

RESULT 7
S55399
genome polyprotein - rabbit hemorrhagic disease virus (isolate BS89)
C:Species: rabbit hemorrhagic disease virus
A:Variety: isolate BS89
C:Date: 15-Jul-1995 #sequence_revision 01-Sep-1995 #text_change 18-Jun-1999
R:Rossi, C.
submitted to the EMBL Data Library, May 1995
A:Reference number: S55399
A:Molecule type: genomic RNA
A:Residues: 1-2344 <ROS>
A:Cross-references: EMBL:X87607; NID:g854640; PIDN:CAA60910.1; PID:g854641
A:Experimental source: isolate BS89
C:Superfamily: rabbit calicivirus RNA-directed RNA polymerase
C:Keywords: polyprotein

Query Match 10.8%; Score 311.5; DB 2; Length 2344;
Best Local Similarity 27.4%; Pred. No. 2.3e-14;
Matches 127; Conservative 60; Mismatches 172; Indels 105; Gaps 20;

QY 17 GAGQLVPEANTAEPISMEPVAGAAAT-----AAATAG-----QV 49
DB 1768 GKARTAPQGEAAGTATTASVPGTTDGLDPGVVATTSSVVTAENSSASIAIATAGIGPPQV 1827
QY 50 NMIDPIMNNYVQAPQGEFTISPNNTPGDILDLQGLPHLNFPLSLAOMYNGWGNMKV 109
DB 1828 DQGETWRTNFY---NDVFTWSVADAPGSLITYVQHSPPNNPFTAVLSQMYAGWAGGMGF 1884
QY 110 KVLLAGNAFTAGKIISCIPPGFAAQNISIA---QATMPPHVIADVRLVEIEVPLEDVR 166
DB 1885 RFIVAGSGVGGRLVAAVIPPG-----IEIGPGLVQRQPHVVIDARSLPEVTIIMPDLR 1939
QY 167 NVLFH-NNDNAPTMRLVCMLYTLP--RASGSSGTDPPFVIAGRVLTCPSDFSLFLVPP 223
DB 1940 PNMVHTGDPGLVPTLVLSVNNLINPFEGSTS-----AIQVTVETRPSEDFEVMIRAP 1994
QY 224 NVEQTKPFSVPNPLNTLSNSRVPSLIKSMVSRDHQGMVOFQ-----NGRVTLD 274
DB 1995 S--SKTVDISPAAGLT-----PVLTVGNDNRNMGQIVGLQVPVPGFSTCNRHNNLN 2046
QY 275 GQLQGTTPTSASOLCKIRGSVFHANGNGYNLTDELSPYHAFESP-----DGSPYHAFESP 326
DB 2047 GSTYGWSSPRFADIDHRRGSASP-GSNATNVLQFWYANAGS----AVDNPISQVADPDGP 2102
```

```

QY 327 DLGECDDHMEASPTQFNTGDIKQINVKQESAFAPHLGTIQADGLSDSVSVNTNMIKLG 386
DB 2103 DMSFVPFPGPGIPAAGWGWGCAIWNNSN-----SGAPNVTTVQA-----YELG 2144
QY 387 WVSVPVSDGHRGDVDPWVTPRYGSTLTLEAAQ-LAPPIYPPGFGGEA 429
DB 2145 FAT-----GAPGNLQ-----TTNTSGAQTVAKSIYAVVTGTA 2177

RESULT 8
S64740
genome polyprotein - rabbit hemorrhagic disease virus (isolate AST/89)
C:Species: rabbit hemorrhagic disease virus
A:Variety: isolate AST/89
C:Date: 12-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 20-Jun-2000
C:Accession: S64740; S46944; S49018; S65012
R:Casais, R.; Martin-Alonso, J.; Boga, J.; Parra, F.
submitted to the EMBL Data Library, May 1995
A:Description: Genomic organization of rabbit hemorrhagic disease virus determined by
A:Reference number: S64740
A:Accession: S64740
A:Molecule type: genomic RNA
A:Residues: 1-2344 <CAS>
A:Cross-references: EMBL:Z49271; NID:g1182032; PIDN:CAA89265.1; PID:g1150552
A:Experimental source: isolate AST/89
R:Boga, J.; Casais, R.; Marin, M.; Martin-Alonso, J.; Carmenes, R.; Prieto, M.; Parra
submitted to the EMBL Data Library, July 1993
A:Description: Molecular cloning, sequence and expression of the capsid protein gene
A:Reference number: S46944
A:Accession: S46944
A:Molecule type: genomic RNA
A:Residues: 1650-2344 <BOG>
A:Cross-references: EMBL:Z24757; NID:g515622; PIDN:CAA80881.1; PID:g515623
A:Experimental source: isolate AST/89
R:Parra, F.; Boga, J.A.; Marin, M.S.; Casais, R.
Virus Res. 27, 219-228, 1993
A:Title: The amino terminal sequence of VP60 from rabbit hemorrhagic disease virus su
A:Reference number: S49018; MUID:93255896; PMID:8488721
A:Accession: S49018
A:Molecule type: genomic RNA
A:Residues: 1650-1796 <PAR>
A:Cross-references: EMBL:Z24757
A:Experimental source: isolate AST/89
A:Accession: S65012
A:Molecule type: protein
A:Residues: 1767-1779; 1875-1877, 'X', 1879-1881; 1936-1938, 'X', 1940-1941 <PAW>
C:Superfamily: rabbit calicivirus RNA-directed RNA polymerase
C:Keywords: blocked amino end; polyprotein

Query Match 10.6%; Score 307.5; DB 2; Length 2344;
Best Local Similarity 26.3%; Pred. No. 4.5e-14;
Matches 149; Conservative 67; Mismatches 212; Indels 139; Gaps 25;

QY 12 MDG-----TSAGQLVPEANTAE-----PISMEPVAGAAATAA-----ATAG----- 47
DB 1766 MEKARTAPQGEAAGTATTASVPGTTDMDPVGVVATTSSVVTAENSSASIAIATAGIGPPQ 1825
QY 48 QVNMDPIMNNYVQAPQGEFTISPNNTPGDILDLQGLPHLNFPLSLAOMYNGWGNM 107
DB 1826 QVDOQETWRTNFY---NDVFTWSVADAPGSLITYVQHSPPNNPFTAVLSQMYAGWAGM 1892
QY 108 KVKVLLAGNAFTAGKIISCIPPGFAAQNISIA---QATMPPHVIADVRLVEIEVPLED 164
DB 1893 QRFRTIVAGIGVGGRLVAAVIPPG-----IEIGPGLVQRQPHVVIDARSLPEVTIIMP 1937
QY 165 VRNVLFH-NNDNAPTMRLVCMLYTLP--RASGSSGTDPPFVIAGRVLTCPSDFSLFLV 221
DB 1938 LRPNVHTGDPGLVPTLVLSVNNLINPFEGSTS-----AIQVTVETRPSEDFEVMIR 1992
QY 222 PPNVEQTKPFSVPNPLNTLSNSRVPSLIKSMVSRDHQGMVOFQ-----NGRVT 272
```

```
Db 1993 APS--SKTVDSISAPAGLLTT-----PVLTVGVNDNRWNGOIVGLQVPVPGGFSFCNRHWN 2044
Qy 273 LDGOLQGTTPTSASOLCKIRGSPHANGGYNLTTEL-----DGSYPHAFESP-----APIG 324
Db 2045 LMGSTYGVSSPFRGIDGHRRGASYP--GNNATNVLPQWYANAGS---AIDNPISQVAPDG 2100
Qy 325 FPDLGECDDHMEASPTTQFNFGDVKOIKQINVKQESAFAPHLGTIQADGLSDVSVNTNMIK 384
Db 2101 FPDMSFVPFNGPGIIPAAAGWGVFGAIWNSN-----SGAPNVTTVQA-----YE 2142
Qy 385 LGWSPVSDGHRGVDWPIVPRYGSTLTEAAQ-LAPPIYPGGEAI-----VFMSDFPI 439
Db 2143 LGFAT-----GAPGNLQP-----TTNTSGQTVAKSIYAVVTGTAQNPAGLFVMAAGVI 2191
Qy 440 AHGNTGLSVPTCTIPOEFVTHFVNEQAPTRGEAALLHYLDPDTHRNLGFEFKLYPEGFWTCV 499
Db 2192 STPSANAITVTPQPDRIVT-----TEGTPAAAPVGVKNTPIMFASV 2232
Qy 500 PNSS-----GTGQOTLPI 512
Db 2233 RRTGDVNATAGSNGTQVGTGSQPLPV 2259

RESULT 9
A:Reference number: A48562; MUID:92410750; PMID:1529644
A:Accession: A48562
A:Molecule type: genomic RNA
A:Residues: 1-702 <NET>
A:Cross-references: GB:M87481; NID:g334882; PIDN:AAAL6217.1; PID:g334884
A:Note: sequence extracted from NCBI backbone (NCBIN:113564, NCBIPI:113565)
C:Superfamily: feline calicivirus coat protein
C:Keywords: capsid protein; coat protein; glycoprotein
F:208,481,493,545/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 10.4%; Score 300.5; DB 1; Length 702;
Best Local Similarity 25.2%; Pred. No. 2.6e-14;
Matches 137; Conservative 66; Mismatches 209; Indels 131; Gaps 25;

Qy 13 DGTSGAGQLVPEANTAEPIISMPEVAGAATAAGOVNMD-PW-IMNNYVQAPQGEFTI 70
Db 154 DGPAGADIVTEQGTVOQVPVAGSALTTLAAASTGKTVDCEWTFPSYHTA----VNW 209
Qy 71 SPNPTPGDILFDLQGLPHNPLSHLAOMYNGWGNMKVKVLLAGNAFTAGKIIISCIPP 130
Db 210 STTEAQQKILFRALSPELNPYLRHISLYSTWSSGIDVRETVSGGVFGKLAALIYPP 269
Qy 131 GFAAONISTAQATMPPHVIADVRVLEPTEVPLEDVRNLFHNDNAPTRMLVCMLYTPL- 189
Db 270 GI--EPVESPTMLQPHVLFDRQTEPVLFITPIDIRKTLYHSMDDTTRLRVIMVYNELI 327
Qy 190 -----RAGSSSGTDPFVIAGRVLTCPSPDFSLFLVPPNVEQKTPFSPVN--LPINT 241
Db 328 NPYEQSEPKSSCSIT-----VETRPSSDFTFSLKPPG--SLLKHSIPSDLI PRNS 377
Qy 242 ---LSNSRVPSLTKSMVSRDHQMVQFQNGR-----VTLD----- 274
Db 378 RHWNGNRWNSSTIDGFV-----QPRVFSQNRHDFDFTTGTWSTPYIPIEVTLEKLDRG 432
Qy 275 GOLQGTTPTSASOL-----CKIRGSVFHANGGYNLTTEL-----DGSYPHAFESPA 321
Db 433 GQYFKVDTDEKSLVPLGCLDQGWPDPTTIPATWASNGNYDTVAEYRITNNGTHFKGFYING 492
Qy 322 PI-----GPPDLGECDDHMEASPTTQFNFGDVKOIKQINVKQESAFAP-----HLGTIOA 369
```

```
Db 493 NLITTKVKGSDNLGET--QOTSRTLFASVG-----NYKDQNTINPTHKITSNSLVVYDA 543
Qy 370 DGLSDSVNVT-----NMIAKLGWV-----SPVSDGHRGVDWPIVPRYGSTLTERAQA--P 419
Db 544 NNYSAATAKTTHWSTSMHSLGYLVLDSPV-----GSDSTKVVRIATLP 587
Qy 420 PIYPPGGEAIVFEMSDFPITAHGTNGLSVPTCTIPOEFVTHFVNEQAPTRGEAALLHY-LD 478
Db 588 EAFTNG--GNPPVFTTKIQTIGH-----FDRHTKCFNSQVLMTSOKLAENHYTLP 636
Qy 479 PDT 481
Db 637 PDS 639

RESULT 10
Qy 34 EPVAGAATAAGOVNMDPWIMNNYVQAPQGEFTI-----SPNPTPGDILFDLQGLP 87
Db 145 EPSAQMTAADMAVTKSVDSSEW-----EAFPSHTSVNWSSTETQCKILFKQSLGP 195
Qy 88 HLNPFSLHQAOMYNGWGNMKVKVLLAGNAFTAGKIIISCIPPGF-AAQNISTAQATMFP 146
Db 196 LLNPLYEHLKSLXYVAMSGSEVREFSISGSGVFGKLAALIYVPPGVDPVQSTMLQ----YP 252
Qy 147 HVIADVRVLEPTEVPLEDVRNLFHNDNAPTRMLVCMLY-----TPLRASGSSSGTDPFV 202
Db 253 HVJFDRQVDVPIFSIFDLRSTLHLPDPTDTTSLVIMVYNDLINPYANDSNSSGC---I 309
Qy 203 IAGRVLTCPSPDFSLFLVPP----- 223
Db 310 VT--VETKPGDPKPHLLKPPGSMLTGHSVPSDLIPKSSSLWIGNRYWSDITDFVVRPFV 367
Qy 224 -----NVEOKTPFSPVNL-PLN-TLSNSRVPSLTKSMVSRDHQMVQFQNGRVTLTD 274
Db 368 FQANRHFNFNETAGSAPRFRPTITITISEKSGKL--GIGVATDY-----IVP 414
Qy 275 GOLQGTTPTSASOLCKIRGSVFHANGGYNLTLDGSPYHAFE-SPAPIGPPDLGECDD- 332
Db 415 GIPDGWPDPTTAEADLTTPAGD-YAITSNGNDIT--TGSEYDSTEVIKNTNFRGMVIGCS 471
Qy 333 ---WHMEASPTTQF-----NTGDVVKQINVKQESAFAPHLGTIOADGLSD 374
Db 472 LQRAWGDKKISNTAFITTAIKEGNKIRPSNTIDMTK-LAVYQDT-----HVG-----EE 519
Qy 375 VSVNTNMIKLGWSPVSDGHRGVDP-----WVTPRYGSTLTERAQAOLAPPIYPGFGGEAI 430
Db 520 VQTSDDALLALGYTGIGEQAIGSDRDRVVRISVUPEVGAR----- 564
Qy 431 VFEMSDFPITAHGTNGLSVPTCTIPOEFVTHFVNEQAPTRGEAALLHY-LDPDT 481
Db 565 IFYKNSIKLGYVIRSIDV-----FNSQILHTSRQLSLNHYLLSPDS 605
```

```
capsid protein - feline calicivirus (strain NADC)
C:Species: feline calicivirus
C:Date: 28-Aug-1985 #sequence_revision 07-Oct-1994 #text_change 20-Sep-1999
C:Accession: JQ2354
R:Seal, B.S.; Riopack, J.F.; Mengeling, W.L.
J. Gen. Virol. 74, 2519-2524, 1993
A:Title: Analysis of feline calicivirus capsid protein genes: identification of varia
A:Reference number: JQ2354; MUID:94065683; PMID:7504075
A:Molecule type: mRNA
A:Residues: -668 <SEA>
A:Cross-references: GB:I09718; NID:g305104; PIDN:AAA16485.1; PID:g305105
C:Superfamily: feline calicivirus coat protein
C:Keywords: capsid protein; coat protein

Query Match 9.6%; Score 277.5; DB 2; Length 668;
Best Local Similarity 23.3%; Pred. No. 1.3e-12;
Matches 124; Conservative 75; Mismatches 178; Indels 155; Gaps 24;

Qy 34 EPVAGAATAAGOVNMDPWIMNNYVQAPQGEFTI-----SPNPTPGDILFDLQGLP 87
Db 145 EPSAQMTAADMAVTKSVDSSEW-----EAFPSHTSVNWSSTETQCKILFKQSLGP 195
Qy 88 HLNPFSLHQAOMYNGWGNMKVKVLLAGNAFTAGKIIISCIPPGF-AAQNISTAQATMFP 146
Db 196 LLNPLYEHLKSLXYVAMSGSEVREFSISGSGVFGKLAALIYVPPGVDPVQSTMLQ----YP 252
Qy 147 HVIADVRVLEPTEVPLEDVRNLFHNDNAPTRMLVCMLY-----TPLRASGSSSGTDPFV 202
Db 253 HVJFDRQVDVPIFSIFDLRSTLHLPDPTDTTSLVIMVYNDLINPYANDSNSSGC---I 309
Qy 203 IAGRVLTCPSPDFSLFLVPP----- 223
Db 310 VT--VETKPGDPKPHLLKPPGSMLTGHSVPSDLIPKSSSLWIGNRYWSDITDFVVRPFV 367
Qy 224 -----NVEOKTPFSPVNL-PLN-TLSNSRVPSLTKSMVSRDHQMVQFQNGRVTLTD 274
Db 368 FQANRHFNFNETAGSAPRFRPTITITISEKSGKL--GIGVATDY-----IVP 414
Qy 275 GOLQGTTPTSASOLCKIRGSVFHANGGYNLTLDGSPYHAFE-SPAPIGPPDLGECDD- 332
Db 415 GIPDGWPDPTTAEADLTTPAGD-YAITSNGNDIT--TGSEYDSTEVIKNTNFRGMVIGCS 471
Qy 333 ---WHMEASPTTQF-----NTGDVVKQINVKQESAFAPHLGTIOADGLSD 374
Db 472 LQRAWGDKKISNTAFITTAIKEGNKIRPSNTIDMTK-LAVYQDT-----HVG-----EE 519
Qy 375 VSVNTNMIKLGWSPVSDGHRGVDP-----WVTPRYGSTLTERAQAOLAPPIYPGFGGEAI 430
Db 520 VQTSDDALLALGYTGIGEQAIGSDRDRVVRISVUPEVGAR----- 564
Qy 431 VFEMSDFPITAHGTNGLSVPTCTIPOEFVTHFVNEQAPTRGEAALLHY-LDPDT 481
Db 565 IFYKNSIKLGYVIRSIDV-----FNSQILHTSRQLSLNHYLLSPDS 605
```

RESULT 11

VCWVFC
coat protein - feline calicivirus (strain F9)
N:Alternate names: capsid protein
C:Species: feline calicivirus
C>Date: 30-Jun-1993 #sequence_revision 30-Jun-1993 #text_change 28-Jul-2000
C:Accession: B43382; C45538; PQ0407; S23702
R:Cartier, M.J.; Milton, I.D.; Meanger, J.; Bennett, M.; Gaskell, R.M.; Turner, P.C.
Virology 190, 443-448, 1992
A:Title: The complete nucleotide sequence of a feline calicivirus.
A:Reference number: A43382; MUID:92410623; PMID:1529544
A:Accession: B43382
A:Molecule type: genomic RNA
A:Residues: 1-671 <CAR2>
A:Cross-references: GB:M86379; NID:g323877; PIDN:AA79327.1; PID:g323879
R:Cartier, M.J.; Milton, I.D.; Turner, P.C.; Meanger, J.; Bennett, M.; Gaskell, R.M.
Arch. Virol. 122, 223-235, 1992
A:Title: Identification and sequence determination of the capsid protein gene of feline calicivirus.
A:Reference number: A45538; MUID:92117861; PMID:1731695
A:Accession: C45538
A:Molecule type: genomic RNA; protein
A:Residues: 1-671 <CAR2>
A:Cross-references: GB:M86379; NID:g323877; PIDN:AA79327.1; PID:g323879
A:Experimental source: strain F9
A:Note: sequence extracted from NCBI backbone (NCBIN:77457, NCBIP:77462)
R:Guiver, M.; Littler, E.; Caul, E.O.; Fox, A.J.
J. Gen. Virol. 73, 2429-2433, 1992
A:Title: The cloning, sequencing and expression of a major antigenic region from the feline calicivirus.
A:Reference number: PQ0407; MUID:93019069; PMID:1402818
A:Accession: PQ0407
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 217-266
A:Cross-references: PIDN:AAB23553.1; PID:g257083
C:Superfamily: feline calicivirus coat protein
C:Keywords: capsid protein; coat protein; glycoprotein
F:177,304,439,459,618/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 9.5%; Score 276; DB 1; Length 671;

Best Local Similarity 26.4%; Pred. NO. 1.6e-12;

Matches 102; Conservative 54; Mismatches 128; Indels 102; Gaps 17;

QY 13 DGTSGAGQLVPEANT-AEPTSMPEVAGATAAATAGQVNMIDPWIMNNYVQAPQGEFTI- 70

Db 127 DGSITA-----PEQTVGVIAPSAQMSAADMATGKSVDSW-----EAFSFFH 173

QY 71 -----SPNTPGDFLDLQGLPFLSHLAQMYNGWGNMKVLLAGNAFTAGKII 125

Db 174 TSNWSTSETQGIKLFQSLGPLLPYLEHLAKLVAVWSGSTEVRFSISGSGVFGKLLA 233

QY 126 SCIPPCGF-AAQNTSIAQATMFPHVADVRVLEPIEVPLEDVNRVLFHNDNAPTMRVLCM 184

Db 234 IVPVPGVDPVQSTMLQ---YPHVLFDARQVEPIFCLPDLKSTLYHLMSTDTTSLVIM 290

QY 185 LY-----TPLRASGSGTDPFVIAGRVLCPSDFSLFLVPPNVQKTFKPSVPLN 240

Db 291 VYNDLIPYANDANSNGC---IVT--VETKPGDFKFKLLKPPG-----SMLTHG 335

QY 241 TLNSRVPSLTKSMVSRDHQGVQFQNGRVTLGDLQGTTPTSASQLKIRGWSFHANG 300

Db 336 SIPLDKPKTSLWIGNRVWSDITDF-----VIRPFVQANR 372

QY 301 GNGYN-----LTELDSGPYHA-----FESPA-PIGFPDL---GEC--- 331

Db 373 HFDENQETAGWTPRFRPSVTITQNGAKLGIGVATDYIVPGIDGWPDTTIPGELIPA 432

QY 332 -DWHM-----EASPTQNTGCVIK 350

Db 433 GDVAITNGTNDITATGTGTADIIK 458

RESULT 12

VCWVFF

coat protein - feline calicivirus (strain CFI/68 FIV)

N:Alternate names: capsid protein

C:Species: feline calicivirus

C>Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 20-Aug-1999

C:Accession: A40507; B40507; T09246

R:Neill, J.D.; Reardon, I.M.; Heinrikson, R.L.

J. Virol. 65, 5440-5447, 1991

A:Title: Nucleotide sequence and expression of the capsid protein gene of feline calicivirus.

A:Reference number: A40507; MUID:91374597; PMID:1716692

A:Accession: A40507

A:Molecule type: genomic RNA

A:Residues: 1-668 <NET>

A:Cross-references: GB:M32819; NID:g323874; PIDN:AAA42925.1; PID:g323875

A:Accession: B40507

A:Molecule type: protein

A:Residues: 373-379; 403-419; 481-489; 560-566 <NE2>

R:Neill, J.D.

submitted to the EMBL Data Library, April 1998

A:Description: Complete nucleotide sequence of feline calicivirus strain CFI/68.

A:Reference number: 216626

A:Accession: T09246

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: genomic RNA

A:Residues: 1-668 <NE3>

A:Cross-references: EMBL:U13992; NID:g3056875; PIDN:AAC13993.1; PID:g537256

A:Experimental source: strain CFI/68 FIV

C:Superfamily: feline calicivirus coat protein

C:Keywords: capsid protein; coat protein; glycoprotein

F:177,301,304,439,459,615/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match

Best Local Similarity 9.3%; Score 269; DB 1; Length 668;

Matches 74; Conservative 39; Mismatches 82; Indels 44; Gaps 9;

```

RESULT 15
JQ2356
capsid protein - feline calicivirus (strain KCD)
C:Species: feline calicivirus
C:Date: 28-Aug-1985 #sequence_revision 07-Oct-1994 #text_change 20-Sep-1999
C:Accession: JQ2356
R:Seal, B.S.; Ridpath, J.F.; Mengeling, W.L.
J. Gen. Virol. 74, 2519-2524, 1993
A:Title: Analysis of feline calicivirus capsid protein genes: identification
A:Reference number: JQ2354; MUID:94065683; PMID:7504075
A:Accession: JQ2356
A:Molecule type: mRNA
A:Residues: 1-668 <SEA>
A:Cross-references: GB:L09719; NID:g305107; PIDN:AAAL6487.1; PID:g305108
C:Superfamily: feline calicivirus coat protein
C:Keywords: capsid protein

Query Match          9.0%   Score 261.5;  DB 2;   Length 668;
Best Local Similarity 31.1%;   Pred. No. 1.9e-11;
Matches 76;   Conservative 37;   Mismatches 86;   Indels 45;   Gaps

```

```

best boundary primary, secondary, tertiary, quaternary, no gaps; Indels *45; Gaps
Matches 76; Conservative 37; Mismatches 86;
QY      19 GOLV-PEANT-AEPISMEPVAGAATAAATAGOVNMIDPWIMNNYVQAPEGFTI----- 70
          |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
Db      128 GSIVTPEQGTVMGVIAEPNAQMSTAADMATKSKVDSEW-----EAFFSFTSVNW 178
          |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
QY      71 SPNNTGCDILFDLGLPHLNPELSHLAOMYNMGVNMKVLLAGNAFTACKIIISCIP 130
          |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
Db      179 STSETGCKILLFKOSLGPLNPYLEHLAKUYAVMSGIEVRFSISGSVGFGKLAAIVVPP 238
          |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
QY     131 GF-AAONISIAQTAMTPHHVIADVRLVEPLEVDNRVNLFHNNNDNAPTMBLVCMLY--- 186
          |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
Db     239 GVDPVQSTSLMQ--YPHVLEDAQOVPFIIFPOLRSTLYHMSDITDTTSILVVMAVNDL 295
          |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
QY     187 -TPLRASGSSGTDPEVIAGRVLTCPSDFSLFLVPPNVNQKTFPFPVNLPLNTLSNS 245
          |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
Db     296 INPYANDSNSGC---IVT--VETPKGSDREHLLKPPG-----SVLTHG 335
          |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
QY     246 RVPS 249
          |||
Db     336 SVPS 339

```

```

DO      300  5918  352
Search completed: March 10, 2003, 19:00:11
Job time : 29 secs

```


GenCore version 5.1.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2003, 18:19:25 ; Search time 38 Seconds
(without alignments)
2955.149 Million cell updates/sec

Title: US-09-926-799-1
Perfect score: 2896
Sequence: 1 MMASKDAPTNDGTSGAGQ.....YQLKPVGTAGPACRLGIRRS 545

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL21.*

- 1: sp.archaea.*
- 2: sp.bacteria.*
- 3: sp.fungi.*
- 4: sp.human.*
- 5: sp.invertebrate.*
- 6: sp.mammal.*
- 7: sp.mhc.*
- 8: sp.organelle.*
- 9: sp.phage.*
- 10: sp.plant.*
- 11: sp.rodent.*
- 12: sp.virus.*
- 13: sp.vertebrate.*
- 14: sp.unclassified.*
- 15: sp.rvirus.*
- 16: sp.bacteriaph.*
- 17: sp.archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2873	99.2	545	12	Q8VA02 human calic
2	2846	98.3	545	12	Q9W182 norwalk-lik
3	2846	98.3	545	12	Q918C8 norwalk-lik
4	2840	98.1	545	12	Q918C2 norwalk-lik
5	2798	96.6	545	12	Q91H12 human calic
6	2744	94.8	545	12	Q91V45 human calic
7	2605	90.0	544	12	Q66A18 desert shie
8	2124	73.3	539	12	Q91V48 human calic
9	1979	68.3	544	12	Q9DU46 chiba virus
10	1975	68.2	544	12	Q91I85 human calic
11	1974	68.2	544	12	Q9QTE7 chiba virus
12	1972.5	68.1	540	12	Q9YS14 norwalk vir
13	1968	68.0	544	12	Q918C5 norwalk-lik
14	1964.5	67.8	546	12	Q91V47 human calic
15	1956.5	67.6	543	12	Q91V43 human calic
16	1956.5	67.6	544	12	Q8VA27 human calic

17	1954	67.5	544	12	Q918D1 norwalk-lik
18	1952	67.4	544	12	Q91EJ2 human calic
19	1946.5	67.2	543	12	Q918B9 norwalk-lik
20	1946	67.2	530	12	Q83884 norwalk vir
21	1942	67.1	544	12	Q91V36 human calic
22	1941	67.0	544	12	Q91V41 human calic
23	1938	66.9	530	12	Q9QV39 norwalk-lik
24	1930.5	66.7	531	12	Q8V735 norwalk-lik
25	1908.5	65.9	517	12	Q91H41 norwalk-lik
26	1900	65.6	530	12	Q83876 norwalk vir
27	1886	65.1	541	12	Q91V42 human calic
28	1274	44.0	522	12	Q8V629 norwalk-lik
29	1268	43.8	522	12	Q8V628 norwalk-lik
30	1260.5	43.5	519	12	Q9YQ22 bovine cali
31	1228	42.4	540	12	Q917X1 norwalk-lik
32	1223	42.2	542	12	Q66293 calicivirus
33	1222	42.2	542	12	Q913B6 human calic
34	1219	42.1	540	12	Q91V50 human calic
35	1215.5	42.0	542	12	Q12305 snow mounta
36	1213	41.9	542	12	Q96877 snow mounta
37	1211	41.8	538	12	Q91I65 human calic
38	1203	41.5	540	12	Q917W8 norwalk-lik
39	1202.5	41.5	535	12	Q915C6 human calic
40	1202.5	41.5	540	12	Q918B0 norwalk-lik
41	1198.5	41.4	548	12	Q88291 small round
42	1196	41.3	537	12	Q913B7 human calic
43	1189	41.1	537	12	Q91H09 human calic
44	1188.5	41.0	548	12	Q917Z5 norwalk-lik
45	1188.5	41.0	548	12	Q916E5 human calic

ALIGNMENTS

RESULT 1

Q8VA02	PRELIMINARY;	PRT;	545 AA.
ID	Q8VA02;		
AC	Q8VA02;		
DT	01-MAR-2002 (TREMBLrel. 20, Created)		
DT	01-MAR-2002 (TREMBLrel. 20, Last sequence update)		
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)		
DE	Capsid protein.		
OS	Human calicivirus NLV/Potsdam 196/2000/DE.		
OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae.		
OC	Norwalk-like viruses.		
OX	NCBI_TaxID=150135;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=HU/NLV/POTSDAM 196/2000/DE;		
RA	Kuenkel U., Schreier E.;		
RT	"Molecular epidemiology of outbreaks of gastroenteritis associated with Norwalk-like viruses in Germany.";		
RL	Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; AF439267; AAL32455.1; ..		
DR	InterPro; IPR004005; Calici_coat.		
DR	Pfam; PF00915; Calici_coat; 1.		
DR	SEQUENCE 545 AA; 58601 MW; 7D47D9F3521D89 CRC64;		

Query Match 99.2%; Score 2873; DB 12; Length 545;
Best Local Similarity 99.1%; Pred. No. 1e-222; 2; Indels 0; Gaps 0;
Matches 540; Conservative 3; Mismatches 0;

QY	1	MMASKDAPTNDGTSGAGOLVPEANTAEPIISMEPVAGATAAATAGQVNMIDPWIMNNY 60	
Db	1	MMASKDAPTNDGTSGAGOLVPEANTAEPIISMEPVAGATAAATAGQVNMIDPWIMNNY 60	
QY	61	VQAPQGEFTISPNPTGDILFDLQGLPHLPFLSHLAQMYNGVGNMKVLLAGNAFTA 120	
Db	61	VQAPQGEFTISPNPTGDILFDLQGLPHLPFLSHLAQMYNGVGNMKVLLAGNAFTA 120	
QY	121	GKIIISICPGFAQNISIAQATMFPHVIADVRVLEPIEVPLEDRVNVLFHNDNAPTMR 180	
Db	121	GKIIISICPGFAQNISIAQATMFPHVIADVRVLEPIEVPLEDRVNVLFHNDNAPTMR 180	

Matches 537; Conservative 2; Mismatches 6; Indels 0; Gaps 0;									
Qy	1	MMASKDAPTNMDGTSAGQLVPEANTAEPI	SMPEPVAGAATAAATAGQVNMIDPWIMNNY	60					
Db	1	MMASKDAPPNMDGTSAGQLVPEANTAEPI	SMPEPVAGAATAAATAGQVNMIDPWIMNNY	60					
Qy	61	VQAPQGEFTTSPNNTPGDILFDLQGLPHLPF	LSHLAQMYNGWGNKVKVLLAGNAFTA	120					
Db	61	VQAPQGEFTTSPNNTPGDILFDLQGLPHLPF	LSHLAQMYNGWGNKVKVLLAGNAFTA	120					
Qy	121	GKIIISCIIPGFAAQNISIAQATMFPHV	IADVRVLEPIEVLDPVNRVLFHNDNAPTMR	180					
Db	121	GKIIISCIIPGFAAQNISIAQATMFPHV	IADVRVLEPIEVLDPVNRVLFHNDNAPTMR	180					
Qy	181	LVCMLYTPLRASGSSGTDPFVITAGRVLT	CPSPDFSLFLVPPNVEQTKPFSPVNLPLN	240					
Db	181	LVCMLYTPLRASGSSGTDPFVITAGRVLT	CPSPDFSLFLVPPNVEQTKPFSPVNLPLN	240					
Qy	241	TLNSRVPSLIKSMVSRDHGMVQFONGRV	TLTGQLOGTTPTSASOLCKIRGSVFHANG	300					
Db	241	TLNSRVPSLIKSMVSRDHGMVQFONGRV	TLTGQLOGTTPTSASOLCKIRGSVFHANG	300					
Qy	301	NGYNLTDELGSPYHAFESPAPIGF	PDLGECDHMEASPTTQFNTGDIKQINVKQESAF	360					
Db	301	NGYNLTDELGSPYHAFESPAPIGF	PDLGECDHMEASPTTQFNTGDIKQINVKQESAF	360					
Qy	361	APHLGTTQADGLSDVSVNTNMIKLGW	SPVSDGHRGVDVPIPRYGSTLTEAAQLAPP	420					
Db	361	APHLGTTQADGLSDVSVNTNMIKLGW	SPVSDGHRGVDVPIPRYGSTLTEAAQLAPP	420					
Qy	421	IYPGFGGAIIVFFMSDFPIAHGTNGL	SVPCPTPOEFVTHFVNEQAPTRGEAALLHYLDPD	480					
Db	421	IYPGFGGAIIVFFMSDFPIAHGTNGL	SVPCPTPOEFVTHFVNEQAPTRGEAALLHYLDPD	480					
Qy	481	THRNLGEBFKLYPEGFMTCVPNSSGT	GPQTLPIINGVVFVSVSRFYQLKPVGTGAPACRL	540					
Db	481	THRNLGEBFKLYPEGFMTCVPNSSGT	GPQTLPIINGVVFVSVSRFYQLKPVGTGAPACRL	540					
Qy	541	GIRRS 545							
Db	541	GIRRS 545							
RESULT 4									
Qy18C2		PRELIMINARY;	PRT;	545 AA.					
AC	Qy18C2								
DT	01-DEC-2001 (TrEMBLrel. 19, Created)								
DT	01-DEC-2001 (TrEMBLrel. 19, Last sequence update)								
DT	01-MAR-2002 (TrEMBLrel. 20, Last annotation update)								
DE	Capsid protein.								
CN	ORF2.								
OS	Norwalk-like virus NLV/Little Rock/316/1994/US.								
OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;								
OC	Norwalk-like viruses.								
OX	NCBI_TaxID=171837;								
RN	[1]								
RP	SEQUENCE FROM N.A.								
RC	STRAIN=HU/NLV/LITTLE ROCK/316/1994/US;								
RX	MEDLINE=97193806; PubMed=9041391;								
RA	Ando T., Monroe S.S., Noel J.S., Glass R.I.;								
RT	"A one-tube method of reverse transcription-PCR to efficiently amplify								
RT	a 3-kilobase region from the RNA polymerase gene to the poly(A) tail								
RT	of small round-structured viruses (Norwalk-like viruses).";								
RL	J. Clin. Microbiol. 35:570-577(1997).								
RN	[2]								
RP	SEQUENCE OF 100-192 FROM N.A.								
RC	STRAIN=HU/NLV/LITTLE ROCK/316/1994/US;								
RX	MEDLINE=98071277; PubMed=9407386;								
RA	Noel J.S., Ando T., Leite J.P., Green K.Y., Dingle K.E., Estes M.K.,								
RA	Seto Y., Monroe S.S., Glass R.I.;								
RT	"Correlation of patient immune responses with genetically								
RT	characterized small round-structured viruses involved in outbreaks of								

RT	nonbacterial acute gastroenteritis in the United States, 1990 to								
RT	1995.";								
RL	J. Med. Virol. 53:372-383(1997).								
RN	[3]								
RP	SEQUENCE FROM N.A.								
RC	STRAIN=HU/NLV/LITTLE ROCK/316/1994/US;								
RX	MEDLINE=20266071; PubMed=10804147;								
RA	Ando T., Noel J.S., Fankhauser R.L.;								
RT	"Genetic classification of 'Norwalk'-like viruses.";								
RL	J. Infect. Dis. 181:S336-S348(2000).								
RN	[4]								
RP	SEQUENCE FROM N.A.								
RC	STRAIN=HU/NLV/LITTLE ROCK/316/1994/US;								
RA	Ando T., Seto Y., Noel J.S., Monroe S.S., Glass R.I., Fankhauser R.L.;								
RL	Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.								
DR	EMBL: AF414405; AAL12968.1; -;								
DR	InterPro: IPR004005; Calici_coat.								
DR	Pfam: PF00915; Calici_coat; 1.								
SQ	SEQUENCE 545 AA; 58666 MW; F55076D61D612E82 CRC64;								
Query Match 98.1%; Score 2840; DB 12; Length 545;									
Best Local Similarity 98.5%; Pred. No. 4.7e-220;									
Matches 537; Conservative 3; Mismatches 5; Indels 0; Gaps 0;									
Qy	1	MMASKDAPTNMDGTSAGQLVPEANTAEPI	SMPEPVAGAATAAATAGQVNMIDPWIMNNY	60					
Db	1	MMASKDAPPNMDGTSAGQLVPEANTAEPI	SMPEPVAGAATAAATAGQVNMIDPWIMNNY	60					
Qy	61	VQAPQGEFTTSPNNTPGDILFDLQGLPHLPF	LSHLAQMYNGWGNKVKVLLAGNAFTA	120					
Db	61	VQAPQGEFTTSPNNTPGDILFDLQGLPHLPF	LSHLAQMYNGWGNKVKVLLAGNAFTA	120					
Qy	121	GKIIISCIIPGFAAQNISIAQATMFPHV	IADVRVLEPIEVLDPVNRVLFHNDNAPTMR	180					
Db	121	GKIIISCIIPGFAAQNISIAQATMFPHV	IADVRVLEPIEVLDPVNRVLFHNDNAPTMR	180					
Qy	181	LVCMLYTPLRASGSSGTDPFVITAGRVLT	CPSPDFSLFLVPPNVEQTKPFSPVNLPLN	240					
Db	181	LVCMLYTPLRASGSSGTDPFVITAGRVLT	CPSPDFSLFLVPPNVEQTKPFSPVNLPLN	240					
Qy	241	TLNSRVPSLIKSMVSRDHGMVQFONGRV	TLTGQLOGTTPTSASOLCKIRGSVFHANG	300					
Db	241	TLNSRVPSLIKSMVSRDHGMVQFONGRV	TLTGQLOGTTPTSASOLCKIRGSVFHANG	300					
Qy	301	NGYNLTDELGSPYHAFESPAPIGF	PDLGECDHMEASPTTQFNTGDIKQINVKQESAF	360					
Db	301	NGYNLTDELGSPYHAFESPAPIGF	PDLGECDHMEASPTTQFNTGDIKQINVKQESAF	360					
Qy	361	APHLGTTQADGLSDVSVNTNMIKLGW	SPVSDGHRGVDVPIPRYGSTLTEAAQLAPP	420					
Db	361	APHLGTTQADGLSDVSVNTNMIKLGW	SPVSDGHRGVDVPIPRYGSTLTEAAQLAPP	420					
Qy	421	IYPGFGGAIIVFFMSDFPIAHGTNGL	SVPCPTPOEFVTHFVNEQAPTRGEAALLHYLDPD	480					
Db	421	IYPGFGGAIIVFFMSDFPIAHGTNGL	SVPCPTPOEFVTHFVNEQAPTRGEAALLHYLDPD	480					
Qy	481	THRNLGEBFKLYPEGFMTCVPNSSGT	GPQTLPIINGVVFVSVSRFYQLKPVGTGAPACRL	540					
Db	481	THRNLGEBFKLYPEGFMTCVPNSSGT	GPQTLPIINGVVFVSVSRFYQLKPVGTGAPACRL	540					
Qy	541	GIRRS 545							
Db	541	GIRRS 545							
RESULT 5									
Qy1HL2		PRELIMINARY;	PRT;	545 AA.					
ID	Qy1HL2								
AC	Qy1HL2;								
DT	01-DEC-2001 (TrEMBLrel. 19, Created)								
DT	01-DEC-2001 (TrEMBLrel. 19, Last sequence update)								
DT	01-MAR-2002 (TrEMBLrel. 20, Last annotation update)								
DE	Capsid.								

OS Human calicivirus NLV/VA98115/1998.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norwalk-like viruses.
OX NCBI_TaxID=165509;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NLV/VA98115/1998;
RA Jiang X., Zhong W.M., Parks T., Huang P.W., Wilton N., Barrett E.,
RA Fulton D., Morrow R., Matson D.O.;
RT "Baculovirus expression and antigenic characterization of the capsid
RT proteins of three Norwalk-like viruses."
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY038598; RAK84673.1; ...
DR InterPro: IPR004005; Calici_coat.
DR Pfam: PF00915; Calici_coat; 1.
SQ SEQUENCE 545 AA; 58295 MW; 258B0C7DEB34DAC CRC64;

Query Match 96.6%; Score 2798; DB 12; Length 545;
Best Local Similarity 96.7%; Pred. No. 1.le-216;
Matches 527; Conservative 7; Mismatches 11; Indels 0; Gaps 0;
QY 1 MMASKDAPTNDGTSGAGQLVPEANTAEPI SMEPVAGAATAAATAGOVNMDPWIMNNY 60
DB 1 MMASKDAPTNDGTSGAGQLVPEANTAEPI SMEPVAGAATAAATAGOVNMDPWIMNNY 60
QY 61 VQAPQGEFTISPNTPGDILFDLQGLPHLNPFLSHLAQMYNGWGMKVKVLLAGNAFTA 120
DB 61 VQAPQGEFTISPNTPGDILFDLQGLPHLNPFLSHLAQMYNGWGMKVKVLLAGNAFTA 120
QY 121 GKIISCIIPPGFAAQNISIAQATMFPVIAVRVLEPIEVLVDVNRVLFHNNONAPTMR 180
DB 121 GKIISCIIPPGFAAQNISIAQATMFPVIAVRVLEPIEVLVDVNRVLFHNNONAPTMR 180
QY 181 LVCMLYTPLRASGSSSGTDPFVIAGRVLTCPSPDFSEFLVPPNVEQTKPFSPNPLN 240
DB 181 LVCMLYTPLRASGSSSGTDPFVIAGRVLTCPSPDFSEFLVPPNVEQTKPFSPNPLN 240
QY 241 TLSNSRVPSLIRSMVSRDHQGVQFQNGRVTLTGQGLGTTPTSAQCKIRGVSFHANG 300
DB 241 TLSNSRVPSLIRSMVSRDHQGVQFQNGRVTLTGQGLGTTPTSAQCKIRGVSFHANG 300
QY 301 GNGYNLTDLGGSPYHAFESPAPIGPFDLGECDWHMEASPTTQFNTGDVVKQINVKQESAF 360
DB 301 GNGYNLTDLGGSPYHAFESPAPIGPFDLGECDWHMEASPTTQFNTGDVVKQINVKQESAF 360
QY 361 APHLGTIQADGLSDVSVNTNMIAGLWSPVSDGHRGVDPWVPIPRYGSTLTEAAQLAPP 420
DB 361 APHLGTIQADGLSDVSVNTNMIAGLWSPVSDGHRGVDPWVPIPRYGSTLTEAAQLAPP 420
QY 421 IYPPGGEATVFFMSDFPIAHGTNGLSVPTTIPQEFVTHFVNEQAPTRGEAALLHYLDPD 480
DB 421 IYPPGGEATVFFMSDFPIAHGTNGLSVPTTIPQEFVTHFVNEQAPTRGEAALLHYLDPD 480
QY 481 THRLNGLFEKLYPEGFTCPVNSSGTGTPQLPVGTVFVSWVSRYQLKPVGTAGPACRL 540
DB 481 THRLNGLFEKLYPEGFTCPVNSSGTGTPQLPVGTVFVSWVSRYQLKPVGTAGPACRL 540
QY 541 GIRRS 545
DB 541 GIRRS 545

RESULT 6
Q91V45 PRELIMINARY; PRT; 545 AA.
ID Q91V45
AC Q91V45;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Capsid protein.
OS Human calicivirus NLV/VA98115/1998.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norwalk-like viruses.

OX NCBI_TaxID=122916;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HU/NLV/BIRMINGHAM/93/UK;
RX MEDLINE=20404883; PubMed=10949950;
RA Green J., Vinje J., Gallimore C.I., Koopmans M., Hale A.D.,
RA Clegg J.C., Chamberlain J., Brown D.W.G.;
RT "Capsid protein diversity among 'Norwalk-like' viruses."
RL Virus Genes 20:227-236(2000).
DR EMBL: AJ277612; CAB89093.1; ...
DR InterPro: IPR004005; Calici_coat.
DR Pfam: PF00915; Calici_coat; 1.
SQ SEQUENCE 545 AA; 58689 MW; DC6EBC6EBECDF7C7 CRC64;
Query Match 94.8%; Score 2744; DB 12; Length 545;
Best Local Similarity 95.8%; Pred. No. 2.5e-212;
Matches 522; Conservative 7; Mismatches 16; Indels 0; Gaps 0;
QY 1 MMASKDAPTNDGTSGAGQLVPEANTAEPI SMEPVAGAATAAATAGOVNMDPWIMNNY 60
DB 1 MMASKDAPTNDGTSGAGQLVPEANTAEPI SMEPVAGAATAAATAGOVNMDPWIMNNY 60
QY 61 VQAPQGEFTISPNTPGDILFDLQGLPHLNPFLSHLAQMYNGWGMKVKVLLAGNAFTA 120
DB 61 VQAPQGEFTISPNTPGDILFDLQGLPHLNPFLSHLAQMYNGWGMKVKVLLAGNAFTA 120
QY 121 GKIISCIIPPGFAAQNISIAQATMFPVIAVRVLEPIEVLVDVNRVLFHNNONAPTMR 180
DB 121 GKIISCIIPPGFAAQNISIAQATMFPVIAVRVLEPIEVLVDVNRVLFHNNONAPTMR 180
QY 181 LVCMLYTPLRASGSSSGTDPFVIAGRVLTCPSPDFSEFLVPPNVEQTKPFSPNPLN 240
DB 181 LVCMLYTPLRASGSSSGTDPFVIAGRVLTCPSPDFSEFLVPPNVEQTKPFSPNPLN 240
QY 241 TLSNSRVPSLIRSMVSRDHQGVQFQNGRVTLTGQGLGTTPTSAQCKIRGVSFHANG 300
DB 241 TLSNSRVPSLIRSMVSRDHQGVQFQNGRVTLTGQGLGTTPTSAQCKIRGVSFHANG 300
QY 301 GNGYNLTDLGGSPYHAFESPAPIGPFDLGECDWHMEASPTTQFNTGDVVKQINVKQESAF 360
DB 301 GNGYNLTDLGGSPYHAFESPAPIGPFDLGECDWHMEASPTTQFNTGDVVKQINVKQESAF 360
QY 361 APHLGTIQADGLSDVSVNTNMIAGLWSPVSDGHRGVDPWVPIPRYGSTLTEAAQLAPP 420
DB 361 APHLGTIQADGLSDVSVNTNMIAGLWSPVSDGHRGVDPWVPIPRYGSTLTEAAQLAPP 420
QY 421 IYPPGGEATVFFMSDFPIAHGTNGLSVPTTIPQEFVTHFVNEQAPTRGEAALLHYLDPD 480
DB 421 IYPPGGEATVFFMSDFPIAHGTNGLSVPTTIPQEFVTHFVNEQAPTRGEAALLHYLDPD 480
QY 481 THRLNGLFEKLYPEGFTCPVNSSGTGTPQLPVGTVFVSWVSRYQLKPVGTAGPACRL 540
DB 481 THRLNGLFEKLYPEGFTCPVNSSGTGTPQLPVGTVFVSWVSRYQLKPVGTAGPACRL 540
QY 541 GIRRS 545
DB 541 GIRRS 545

RESULT 7
Q66418 PRELIMINARY; PRT; 544 AA.
ID Q66418
AC Q66418;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Capsid protein.
OS Desert Shield virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norwalk-like viruses.
OX NCBI_TaxID=33755;
RN [1]
RP SEQUENCE FROM N.A.

```
RC STRAIN_DSV395;
RA MEDLINE=94174735; PubMed=8128635;
RA Lew J.F., Kapikian A.Z., Jiang X., Estes M.K., Green K.Y.;
RT "Molecular characterization and expression of the capsid protein of a
RT Norwalk-like virus recovered from a Desert Shield troop with
RT gastroenteritis.";
RL Virology 200:319-325(1994).
DR EMBL: U0469; AAA16285.1; -.
DR InterPro: IPR004005; Calici.coat.
DR Pfam: PF00915; Calici.coat; 1.
SQ SEQUENCE 544 AA; 58606 MW; 07A4AE7E4BCE81E8 CRC64;

Query Match 90.0%; Score 2605.5; DB 12; Length 544;
Best Local Similarity 89.2%; Pred. No. 3.5e-201;
Matches 486; Conservative 24; Mismatches 34; Indels 1; Gaps 1;

QY 1 MMASKDAPTNMDGTSAGQLVPEANTAEPISMPEVAGATAATAAGOVNMDPWIIMNNY 60
DB 1 MMASKDAPTNMDGTSAGQLVPEANTAEPISMPEVAGATAATAAGOVNMDPWIIMNNY 60
QY 61 VQAPQGEFTISPNTPGDILFDLQGLPHLPFLSHLAQMYNGWGNMKVLLAGNAFTA 120
DB 61 VQAPQGEFTISPNTPGDILFDLQGLPHLPFLSHLAQMYNGWGNMKVLLAGNAFTA 120
QY 121 GKIIISCIPGFAAQNISIAQATMFPHVIADRVLEPIEVLDPVFNLFHNDNAPTMR 180
DB 121 GKIIISCIPGFAAQNISIAQATMFPHVIADRVLEPIEVLDPVFNLFHNDNAPTMR 180
QY 181 LVCMLYTPLRASGSSGTDPFVIAGRVLTCPSPDFSEFLVPPNVEQTKPFSPVPLN 240
DB 181 LVCMLYTPLRASGSSGTDPFVIAGRVLTCPSPDFSEFLVPPNVEQTKPFSPVPLN 240
QY 241 TLSNRVPSLIKSMVSRDHQMVQFQNGRVTLDGQLQGTTPTSASOLCKIRGSVFHANG 300
DB 241 TLSNRVPSLIKSMVSRDHQMVQFQNGRVTLDGQLQGTTPTSASOLCKIRGSVFHANG 300
QY 301 GNGYNLTDELGSPYHAFESPAPIGFDPDLGECDDWHM--EASPTTFQNTGDIKQINVKQESAF 360
DB 301 GNGYNLTDELGSPYHAFESPAPIGFDPDLGECDDWHM--EASPTTFQNTGDIKQINVKQESAF 360
QY 361 APHLGTTQADGLSDVSVNTNMIKLGWSPVSDGHRGDDVPWIIPRYGSLTTEAAQLAPP 420
DB 361 APHLGTTQADGLSDVSVNTNMIKLGWSPVSDGHRGDDVPWIIPRYGSLTTEAAQLAPP 420
QY 421 IYPGFGGEAIVFFMSDFPIAHGTNGLSVPCPTIPOEFVTHFVNEQAPTRGEAALLHYLD 480
DB 421 IYPGFGGEAIVFFMSDFPIAHGTNGLSVPCPTIPOEFVTHFVNEQAPTRGEAALLHYLD 480
QY 480 IYPGFGGEAIVFFMSDFPIAHGTNGLSVPCPTIPOEFVTHFVNEQAPTRGEAALLHYLD 479
DB 480 IYPGFGGEAIVFFMSDFPIAHGTNGLSVPCPTIPOEFVTHFVNEQAPTRGEAALLHYLD 479
QY 481 THRNGLGFEKLYPEGFMTCVPNSSGTGPTLPINGVVFVSWSRFYQLKPVGTAGPACRL 540
DB 481 THRNGLGFEKLYPEGFMTCVPNSSGTGPTLPINGVVFVSWSRFYQLKPVGTAGPACRL 540
QY 541 GIRRS 545
DB 540 GIRRS 544

RESULT 8
Q91V48 PRELIMINARY; PRT; 539 AA.
AC Q91V48;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Capsid protein.
OS Human calicivirus HU/NLV/Winchester/94/UK.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norwalk-like viruses.
OX NCBI_TaxID=122913;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HU/NLV/WINCHESTER/94/UK;
RA MEDLINE=20404863; PubMed=10949950;
RX
```

```
RA Green J., Vinje J., Gallimore C.I., Koopmans M., Hale A.D.,
RA Clegg J.C., Chamberlain J., Brown D.W.G.;
RT "Capsid protein diversity among 'Norwalk-like' viruses.";
RL Virus Genes 20:227-236(2000).
DR EMBL: AJ277609; CAB89090.1; -.
DR InterPro: IPR004005; Calici.coat.
DR Pfam: PF00915; Calici.coat; 1.
SQ SEQUENCE 539 AA; 58120 MW; 14E612D7A576395D CRC64;

Query Match 73.3%; Score 2124; DB 12; Length 539;
Best Local Similarity 72.7%; Pred. No. 1.9e-162;
Matches 397; Conservative 56; Mismatches 83; Indels 10; Gaps 6;

QY 1 MMASKDAPTNMDGTSAGQLVPEANTAEPISMPEVAGATAATAAGOVNMDPWIIMNNY 60
DB 1 MMASKDAPTNMDGTSAGQLVPEANTAEPISMPEVAGATAATAAGOVNMDPWIIMNNY 60
QY 61 VQAPQGEFTISPNTPGDILFDLQGLPHLPFLSHLAQMYNGWGNMKVLLAGNAFTA 120
DB 61 VQAPQGEFTISPNTPGDILFDLQGLPHLPFLSHLAQMYNGWGNMKVLLAGNAFTA 120
QY 121 GKIIISCIPGFAAQNISIAQATMFPHVIADRVLEPIEVLDPVFNLFHNDNAPTMR 180
DB 121 GKIIISCIPGFAAQNISIAQATMFPHVIADRVLEPIEVLDPVFNLFHNDNAPTMR 180
QY 181 LVCMLYTPLRASGSSGTDPFVIAGRVLTCPSPDFSEFLVPPNVEQTKPFSPVPLN 240
DB 181 LVCMLYTPLRASGSSGTDPFVIAGRVLTCPSPDFSEFLVPPNVEQTKPFSPVPLN 240
QY 241 TLSNRVPSLIKSMVSRDHQMVQFQNGRVTLDGQLQGTTPTSASOLCKIRGSVFHANG 300
DB 241 TLSNRVPSLIKSMVSRDHQMVQFQNGRVTLDGQLQGTTPTSASOLCKIRGSVFHANG 300
QY 301 GNGYNLTDELGSPYHAFESPAPIGFDPDLGECDDWHM--EASPTTFQNTGDIKQINVKQES 358
DB 301 GNGYNLTDELGSPYHAFESPAPIGFDPDLGECDDWHM--EASPTTFQNTGDIKQINVKQES 358
QY 359 AFAPHLGTTQADGLSDVSVNTNMIKLGWSPVSDGHRGDDVPWIIPRYGSLTTEAAQLA 418
DB 359 AFAPHLGTTQADGLSDVSVNTNMIKLGWSPVSDGHRGDDVPWIIPRYGSLTTEAAQLA 418
QY 419 PIYPGFGGEAIVFFMSDFPIAHGTNGLSVPCPTIPOEFVTHFVNEQAPTRGEAALLHYLD 478
DB 419 PIYPGFGGEAIVFFMSDFPIAHGTNGLSVPCPTIPOEFVTHFVNEQAPTRGEAALLHYLD 478
QY 479 PDTHRNGLGFEKLYPEGFMTCVPNSSGTGPTLPINGVVFVSWSRFYQLKPVGTAGPAC 538
DB 479 PDTHRNGLGFEKLYPEGFMTCVPNSSGTGPTLPINGVVFVSWSRFYQLKPVGTAGPAC 538
QY 539 RL3IRR 544
DB 533 RL3VRR 538

RESULT 9
Q9DU46 PRELIMINARY; PRT; 544 AA.
AC Q9DU46;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE Capsid protein.
OS Chiba virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norwalk-like viruses.
OX NCBI_TaxID=99565;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HU/NLV/CHIBA 407/1987/JP;
RA Someya T., Takeda N., Miyamura T.;
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
```

```
RC STRAIN=HU/NLV/CHIBA 407/1987/JIP;
RX MEDLINE=20569531; PubMed=11118371;
RA Someya Y., Takeda N., Miyamura T.;
RT "Complete nucleotide sequence of the Chiba virus genome and functional
RL expression of the 3C-like protease in Escherichia coli.";
RL Virology 278:490-500(2000).
DR EMBL; AB042808; BAB18267.1; -.
DR InterPro; IPR004005; Calici_Coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 544 AA; 58351 MW; 1DB7A4E81FBB246F CRC64;

Query Match      68.3%; Score 1979; DB 12; Length 544;
Best Local Similarity 66.8%; Pred. No. 9e-151;
Matches 368; Conservative 68; Mismatches 101; Indels 14; Gaps 6;

Qy 1 MMASKDAPTNDGTSAGQLVPEANTAEPISEMPVAGAAATAAGQVNMIDPWIMNNY 60
Db 1 MMASKDAPTNDGTSAGQLVPEANTAEPISEMPVAGAAATAAGQVNMIDPWIMNNY 60
Qy 61 VQAPQGEFTISPNTPGDILFDLQGLPHLNPFLSHLAQMYNGWGNMKVYLLAGNAFTA 120
Db 61 VQAPQGEFTISPNTPGDILFDLQGLPHLNPFLSHLAQMYNGWGNMKVYLLAGNAFTA 120
Qy 61 VQAPQGEFTISPNTPGDILFDLQGLPHLNPFLSHLAQMYNGWGNMKVYLLAGNAFTA 120
Db 61 VQAPQGEFTISPNTPGDILFDLQGLPHLNPFLSHLAQMYNGWGNMKVYLLAGNAFTA 120
Qy 121 GKIIISCIIPPGFAAQNISIAQATMFPHVIADVRVLEPIEVLDPVLEDRVNLHNDNAPTMR 180
Db 121 GKIIISCIIPPGFAAQNISIAQATMFPHVIADVRVLEPIEVLDPVLEDRVNLHNDNAPTMR 180
Qy 121 GKVIICVPPGFGQSRRTLSIAQATLFPVIAVRLTDPVEVPLEDRVNLHNDTQPTMR 180
Db 121 GKVIICVPPGFGQSRRTLSIAQATLFPVIAVRLTDPVEVPLEDRVNLHNDTQPTMR 180
Qy 181 LVCMLYTPLRASGSSGTDPPFVIAGRVLTCPSPDFSLFLVPPNVEQTKPFSPVNLPLN 240
Db 181 LVCMLYTPLRASGSSGTDPPFVIAGRVLTCPSPDFSLFLVPPNVEQTKPFSPVNLPLN 240
Qy 181 LVCMLYTPLRASGSSGTDPPFVIAGRVLTCPSPDFSLFLVPPNVEQTKPFSPVNLPLN 240
Db 181 LVCMLYTPLRASGSSGTDPPFVIAGRVLTCPSPDFSLFLVPPNVEQTKPFSPVNLPLN 240
Qy 241 TLSNRVPSLIKSMVSRDHQWQFQNGRVTLDSQLOGCTTPTASOLCKIRGSVFHANG 300
Db 241 TLSNRVPSLIKSMVSRDHQWQFQNGRVTLDSQLOGCTTPTASOLCKIRGSVFHANG 300
Qy 241 YLSNSRIPNPIEGMSLSPDQTONVQFQNGRCTIDGQPLGTPVSVSOLCKFRGRI--TSG 298
Db 241 YLSNSRIPNPIEGMSLSPDQTONVQFQNGRCTIDGQPLGTPVSVSOLCKFRGRI--TSG 298
Qy 301 GNGYNLTDELDSGYHAFESPAPIGFDPDLGECWHEAS--PTTQFNTGDVVIKQINVKQES 358
Db 301 GNGYNLTDELDSGYHAFESPAPIGFDPDLGECWHEAS--PTTQFNTGDVVIKQINVKQES 358
Qy 359 QRVNLTELDSGSPKAAAPAPAGFPDLGSCDWHIEMSKIPNSSTQNPNIVTSVVPNSQ 358
Db 359 QRVNLTELDSGSPKAAAPAPAGFPDLGSCDWHIEMSKIPNSSTQNPNIVTSVVPNSQ 358
Qy 359 AFAPHLGTIQADGLSDVSNTNMIAGLVWSPVSDGHRGDVDPWIPRYGSTLITEAAQLA 418
Db 359 AFAPHLGTIQADGLSDVSNTNMIAGLVWSPVSDGHRGDVDPWIPRYGSTLITEAAQLA 418
Qy 359 QFVPHLSLITLD--ENVSSGGDYIGTIQWTSPPSDSGGANTFNWKIPDYGSSLAESQLA 416
Db 359 QFVPHLSLITLD--ENVSSGGDYIGTIQWTSPPSDSGGANTFNWKIPDYGSSLAESQLA 416
Qy 419 PPIYPGFGGEAIVFFMSDFPIAHGTNGLS---VPCTIPEQEVTHFVNEQAPTRGEAALL 474
Db 419 PPIYPGFGGEAIVFFMSDFPIAHGTNGLS---VPCTIPEQEVTHFVNEQAPTRGEAALL 474
Qy 417 PAVYPPGFENEVIFYMASIP---GPNQSGSPNLVPCLLPQEVITHFISEQAPIQGEAALL 473
Db 417 PAVYPPGFENEVIFYMASIP---GPNQSGSPNLVPCLLPQEVITHFISEQAPIQGEAALL 473
Qy 475 HYLDPDTHRNLFGEKLYPEGFMTCVPNNSGTPQTLPIGVFVSVWSRYQLKPVGTA 534
Db 475 HYLDPDTHRNLFGEKLYPEGFMTCVPNNSGTPQTLPIGVFVSVWSRYQLKPVGTA 534
Qy 474 HYVDPDTNRNLGFEKLYPGGYLTCPVNSSTGPPQLPLDGVFVFASWVSREYQLKPVGTA 533
Db 474 HYVDPDTNRNLGFEKLYPGGYLTCPVNSSTGPPQLPLDGVFVFASWVSREYQLKPVGTA 533
Qy 535 GPA-CRLGIRR 544
Db 535 GPA-CRLGIRR 544
Qy 534 GPARGRLGVRR 544
Db 534 GPARGRLGVRR 544
```

```
RESULT 10
Q91185 ID Q91185 PRELIMINARY; PRT; 544 AA.
AC Q91185;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Capsid protein.
OS Human calicivirus HU/NLV/Koblentz/433/2000/DE.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norwalk-like viruses.
OC Norwalk-like viruses.
OX NCBI_TaxID=165252;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HU/NLV/KOBLENZ 433/2000/DE;
RA Kuenkel U., Schreier E.;
RT "Molecular epidemiology of outbreaks of gastroenteritis associated
```

```
RT with Norwalk-like viruses in Germany.";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF394960; AAK72048.1; -.
DR InterPro; IPR004005; Calici_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 544 AA; 58222 MW; 1B39A595B733A6A8 CRC64;

Query Match      68.2%; Score 1975; DB 12; Length 544;
Best Local Similarity 66.8%; Pred. No. 1.9e-150;
Matches 368; Conservative 68; Mismatches 101; Indels 14; Gaps 6;

Qy 1 MMASKDAPTNDGTSAGQLVPEANTAEPISEMPVAGAAATAAGQVNMIDPWIMNNY 60
Db 1 MMASKDAPTNDGTSAGQLVPEANTAEPISEMPVAGAAATAAGQVNMIDPWIMNNY 60
Qy 61 VQAPQGEFTISPNTPGDILFDLQGLPHLNPFLSHLAQMYNGWGNMKVYLLAGNAFTA 120
Db 61 VQAPQGEFTISPNTPGDILFDLQGLPHLNPFLSHLAQMYNGWGNMKVYLLAGNAFTA 120
Qy 121 GKIIISCIIPPGFAAQNISIAQATMFPHVIADVRVLEPIEVLDPVLEDRVNLHNDNAPTMR 180
Db 121 GKIIISCIIPPGFAAQNISIAQATMFPHVIADVRVLEPIEVLDPVLEDRVNLHNDNAPTMR 180
Qy 181 LVCMLYTPLRASGSSGTDPPFVIAGRVLTCPSPDFSLFLVPPNVEQTKPFSPVNLPLN 240
Db 181 LVCMLYTPLRASGSSGTDPPFVIAGRVLTCPSPDFSLFLVPPNVEQTKPFSPVNLPLN 240
Qy 181 LVCMLYTPLRASGSSGTDPPFVIAGRVLTCPSPDFSLFLVPPNVEQTKPFSPVNLPLN 240
Db 181 LVCMLYTPLRASGSSGTDPPFVIAGRVLTCPSPDFSLFLVPPNVEQTKPFSPVNLPLN 240
Qy 241 TLSNRVPSLIKSMVSRDHQWQFQNGRVTLDSQLOGCTTPTASOLCKIRGSVFHANG 300
Db 241 TLSNRVPSLIKSMVSRDHQWQFQNGRVTLDSQLOGCTTPTASOLCKIRGSVFHANG 300
Qy 241 YLSNSRIPNPIEGMSLSPDQTONVQFQNGRCTIDGQPLGTPVSVSOLCKFRGRI--TSG 298
Db 241 YLSNSRIPNPIEGMSLSPDQTONVQFQNGRCTIDGQPLGTPVSVSOLCKFRGRI--TSG 298
Qy 301 GNGYNLTDELDSGYHAFESPAPIGFDPDLGECWHEAS--PTTQFNTGDVVIKQINVKQES 358
Db 301 GNGYNLTDELDSGYHAFESPAPIGFDPDLGECWHEAS--PTTQFNTGDVVIKQINVKQES 358
Qy 359 QRVNLTELDSGSPKAAAPAPAGFPDLGSCDWHIEMSKIPNSSTQNPNIVTSVVPNSQ 358
Db 359 QRVNLTELDSGSPKAAAPAPAGFPDLGSCDWHIEMSKIPNSSTQNPNIVTSVVPNSQ 358
Qy 359 AFAPHLGTIQADGLSDVSNTNMIAGLVWSPVSDGHRGDVDPWIPRYGSTLITEAAQLA 418
Db 359 AFAPHLGTIQADGLSDVSNTNMIAGLVWSPVSDGHRGDVDPWIPRYGSTLITEAAQLA 418
Qy 359 QFVPHLSLITLD--ENVSSGGDYIGTIQWTSPPSDSGGANTFNWKIPDYGSSLAESQLA 416
Db 359 QFVPHLSLITLD--ENVSSGGDYIGTIQWTSPPSDSGGANTFNWKIPDYGSSLAESQLA 416
Qy 419 PPIYPGFGGEAIVFFMSDFPIAHGTNGLS---VPCTIPEQEVTHFVNEQAPTRGEAALL 474
Db 419 PPIYPGFGGEAIVFFMSDFPIAHGTNGLS---VPCTIPEQEVTHFVNEQAPTRGEAALL 474
Qy 417 PAVYPPGFENEVIFYMASIP---GPNQSGSPNLVPCLLPQEVITHFISEQAPIQGEAALL 473
Db 417 PAVYPPGFENEVIFYMASIP---GPNQSGSPNLVPCLLPQEVITHFISEQAPIQGEAALL 473
Qy 475 HYLDPDTHRNLFGEKLYPEGFMTCVPNNSGTPQTLPIGVFVSVWSRYQLKPVGTA 534
Db 475 HYLDPDTHRNLFGEKLYPEGFMTCVPNNSGTPQTLPIGVFVSVWSRYQLKPVGTA 534
Qy 474 HYVDPDTNRNLGFEKLYPGGYLTCPVNSSTGPPQLPLDGVFVFASWVSREYQLKPVGTA 533
Db 474 HYVDPDTNRNLGFEKLYPGGYLTCPVNSSTGPPQLPLDGVFVFASWVSREYQLKPVGTA 533
Qy 535 GPA-CRLGIRR 544
Db 535 GPA-CRLGIRR 544
Qy 534 GPARGRLGVRR 544
Db 534 GPARGRLGVRR 544
```

```
RESULT 11
Q9QTE7 ID Q9QTE7 PRELIMINARY; PRT; 544 AA.
AC Q9QTE7;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE Capsid protein.
OS Chiba virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norwalk-like viruses.
OC Norwalk-like viruses.
OX NCBI_TaxID=99565;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CHIBA 407;
RA Natori K., Suzuki K., Yamakawa Y., Tatsumi M., Sakae K., Kobayashi S.,
RA Shinozaki K., Ishiko H., Miyamura T., Takeda N.;
RT "Expression and self-assembly of capsid proteins of the Chiba virus, a
RL genetically distinct Norwalk-like virus.";
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB022679; BAA82106.1; -.
```

DR	InterPro: IPR004005; Calici_coat.
DR	Pham: PF00915; Calici_coat; 1.
SQ	SEQUENCE 544 AA; 58323 MW; ABD1CFC4F93D872 CRC64;
Query Match	
Best Local Similarity 68.2%; Score 1974; DB 12; Length 544;	
Matches 367; Conservative 68; Mismatches 102; Indels 14; Gaps	
Qy	1 MMASKDAPTINMDCTSGAGOLVPEANTAEPISEMPVAGATAAATAAGOVNMIDPWIMNNY 60
Db	1 MMASKDAPESADGATGAGOLVPEVNTADPIPDPVAGSSTALATAGQVNLIDPWIIINF 60
Qy	61 VQAPQGEFTISPNTFGDILDLQGLPHLNPFLSHLAQMYNGHWGNKVKVLLAGNAFTA 120
Db	61 VQAPQGEFTISPNTFGDVLVLDLQGLPHLNPFLSHLSQMYNGHWGNMVRVVLGNAFTA 120
Qy	121 GKIIISICPGFAAQNISIAQATMFPHVITADVRVLEPIEVLDDVNRVLFHNNDNAPTWR 180
Db	121 GKVIICVPPGFSRSLISIAQATLPHVITADVTLDPEVLEDDVNRVLYHNNDTQPTWR 180
Qy	181 LVCMLYPLRASGSSGTSDFVIAGRVLTCTPSPDFSLFLVLPENVEOKTKPFSVPLNPLN 240
Db	181 LLCMLYPLRTGGASGCTDSFVVAGRVLTCPGPDFNFLVLPVTVQKTRPFTVPNPLK 240
Qy	241 TLSNRPVSLIKSMWSRDHGMQVQNGRVTLGDLOGLTTPTSASQLCKIRGSVFHANG 300
Db	241 YLSNRPINPIEGMSLSPDOTNVQFQNGRCTIDGQPLGTPVSVSQLCKFRGRI--TSG 298
Qy	301 GNGVNLTELDGSPVHAPESAPIGPDLGCDWHMEAS--PTQFNTGQDVQIKQINVKQES 358
Db	299 QRVLNLTELDGSPFMGAPAGPDLGSCDWHIEMSKIPNSSTQNNPIVTVNSVKPNQSO 358
Qy	359 AFAPHLGTIQADGLSDVSVNTMIKLGWSPVSDGHRGVDVDPWIPRYGSTLTTEAAQLA 418
Db	359 QFVPHLSITLD--ENVSSGGDYIGTIQWTSPPSDSGGANTFNWKIPDYGSSLAESAQLA 416
Qy	419 PPIYPGPGGEAIVFFMSDFPIAHGTNGLS----VPCTIPEQFVTHFVNQAPTRGGAALL 474
Db	417 PAVYPGPFNEVIVFMASIP---GPNQSGSPNLVPCLLPQEIYTHFISEQAPIQGEAALL 473
Qy	475 HYLPDPDTHNLGEBKLYPEGFWTCVPNSSGTGTLPINGCVFVFSWVSREFYOLKPVGTA 534
Db	474 HYVPDPTNRNLGEBKLYPGGYLCVTPNSSSTGTGQOPLDGVFVFASWVSREFYOLKPVGTA 533
Qy	535 GPA-CRLGIIR 544
Db	534 GPARGRLGVRR 544
RESULT 12	
QyVS14	PRELIMINARY; PRT; 540 AA.
ID	QYVS14
AC	QYVS14
DT	01-MAY-1999 (TrEMBLrel. 10, Created)
DT	01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT	01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE	Capsid protein.
OS	Norwalk virus.
OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC	Norwalk-like viruses.
OX	NCBI_TaxID:11983;
RN	{1}
RN	SEQUENCE FROM N.A.
RC	STRAIN-NORMAL-LIKE;
RC	MEINLONE-20254531; Pubmed=10795514;
RA	Schreier E., Doering F., Kuenkel U.;
RT	"Molecular epidemiology of outbreaks of gastroenteritis associated
RT	with small round structured viruses in Germany in 1997/98.";
RL	Arch. Virol. 145:443-453(2000).
RN	[2]
RP	SEQUENCE FROM N.A.
RC	STRAIN-NORMAL-LIKE;
RA	Schreier E., Doering F., Kuenkel U.;

Query Match	68.1%	Score 1972.5	DB 12	Length 540
Best Local Similarity	67.7%	Pred. No. 3e-150		
Matches 377	Conservative 60	Mismatches 89	Indels 31	Gaps
Qy 1	MMWASKDAPNTMDGTSGAGOLVPEANTAP	ISMPEVAGAAATAAGAGVNMIDPIMNNY	60	
Db 1	MMWASKDAPPSDCASAGOLVPEVNTAQ	ISMNDPVAGASTAVATAGVNMIDPWFNNF	60	
Qy 61	VQAPQGEFTISPNTPGDILDLQGLPHLNFL	SHLAQMYNGWGMKVKVLLAGNAFTA	120	
Db 61	VQAPQGEFTISPNTPGDILDLQGLPHLNFL	LAHLSQMYNGWGMVNRVILLAGNAFTA	120	
Qy 121	GKILISCIPIGCFAAQNISATAQMTFPHV	ADRVLYLEPIEVLVEDVNRVLFHNNDN	179	
Db 121	GKVIICVPGFDFARILITTAQATLFFHL	LADVRLTEPVELPVEDVNRVLYHNSSQ	180	
Qy 180	RLVCLMYTPLRASGSSGGTDPFVIAGRV	LTCPSDFSFLVPPNVEQKTKPFSVNLPL	239	
Db 181	RLVAMLYTPLRTCGGSGGTDAFVVAAGR	VLTCAPDFSFLVPPSVEQKTRFVSVPNPL	240	
Qy 240	NTLSNSRVSLIKSMVSRDHGQWQFONGR	VRTLDGQLGTTPTTSASQLCKIRGSVFHAN	299	
Db 241	KDLSNSRVPTLIOGMFVSPDVNGSVQF	ONGRCQIDCQLGTTPTVLSQLCKIRGRT	--SS 298	
Qy 300	GGNGYNITELDGGSPYIAFESPAIGPD	PLGECQWHM-----EASPTTQNTGDVTK	350	
Db 299	NTRVLNLSEVDGTFPVPLESAPVGP	DPDIGGCDWHVGTFEARDQDPSONVTFATND	--- 355	
Qy 351	QINVKQESAFAPHLGTI--	QAQGLSDSVNTNMIAKLGWVSPVSDG	HGCDVDPWVPIRY 407	
Db 356	-----SSEVPYLGISPHNGDGFH	-----SGDIIGSLDWISAPSDGSALDV	--WSIPKY 402	
Qy 408	GSTLTEAAQLAPPTYPGFGCAIVFMS	DFPIAHGTNGLSVPCTIPOEFTVTHFVNEQAPT	467	
Db 403	GSSLPDVTHLAPVFPFGFGEVILYFHS	KFPGSGPTDKLRVCLPIPOEITHFCNEQAPI	462	
Qy 468	RGEAALLHYLDPTHRNGLBFEKLYL	PEGMTFCVPSNGSGTQPTLPINGVFVFSW	VSRRFFQV 527	
Db 463	AGEAALLHYVDPTGRNLBFEKLYL	PDGEMTCVPNSISGSGTQLPINGVFVFSW	SRRFQV 522	
Qy 528	LKPVGTAGPACRLGIRR	544		
Db 523	LKPVGTAASAARRUGLRR	539		
RESULT 13				
Q918C5	PRELIMINARY;	PRG;	544	AA.
ID Q918C5	PRELIMINARY;	PRG;	544	AA.
AC Q918C5	Q918C5			
DT 01-DEC-2001	(TtEMBLrel. 19, Created)			
DT 01-DEC-2001	(TtEMBLrel. 19, Last sequence update)			
DT 01-MAR-2002	(TtEMBLrel. 20, Last annotation update)			
DE Capsid protein.				
GN ORF2.				
OS Norwalk-like virus NLV/Baltimore/277/1993/US.				
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;				
OC Norwalk-like viruses.				
OX NCBI_TaxID=171836;				
RN [1]				
RP SEQUENCE FROM N.A.				
RC STRAIN=HU/NLV/BALTIMORE/277/1993/US;				
RC MEDLINE=97193806; PubMed=9041391;				
RA Ando T., Monroe S.S., Noel J.S., Glass R.I.;				
RT "A one-tube method of reverse transcription-PCR to efficiently amplify				
RT a 3-kilobase region from the RNA polymerase gene to the poly(A) tail				
RT of small round-structured viruses (Norwalk-like viruses).";				
RL J. Clin. Microbiol. 35:570-577(1997).				

AC	Q91V43;	
DT	01-OCT-2000 (TrEMBLrel. 15, Created)	
DT	01-OCT-2000 (TrEMBLrel. 15, Last sequence update)	
DT	01-DEC-2001 (TrEMBLrel. 19, Last annotation update)	
DE	Capsid protein.	
OS	Human calicivirus HU/NLV/Musgrove/89/UK.	
OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;	
OC	Norwalk-like viruses.	
OX	NCBI_TaxID=122918;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN-HU/NLV/MUSGROVE/89/UK;	
RX	MEDLINE=20404883; PubMed=10949950;	
RA	Green J., Vinje J., Gallimore C.I., Koopmans M., Hale A.D.,	
RA	Clegg J.C., Chamberlain J., Brown D.W.G.;	
RT	"Capsid protein diversity among 'Norwalk-like' viruses.";	
RL	Virus Genes 20:227-236(2000).	
DR	EMBL: AJ277614; CAB89095.1; -	
DR	InterPro: IPR004005; Calici_coat.	
DR	Pfam: PF00915; Calici_coat; 1.	
SQ	SEQUENCE 543 AA; 58216 MW; 11F97A3989CCC540 CRC64;	
Query Match 67.6%; Score 1956.5; DB 12; Length 543;		
Best Local Similarity 67.8%; Pred. No. 5.8e-149;		
Matches 377; Conservative 65; Mismatches 189; Indels 25; Caps 10;		
Qy	1 MMASKDAPTNMDGTSGAGQLVPEANTAEPTSMPEVAGATAATAAGVNMIDPWIMNNY 60	
Db	1 MMASKDATPSADGANGAGOLVPEVNNAXPLPLDPVAGASTALATAGVNMIDPWIFNF 60	
Qy	61 VQAPQGETTSPNTPGDIILFDLQGLPHLNFILSHLAQMYNGWGNKVKVLLAGNAFTA 120	
Db	61 VQAPQGETTSPNTPGDIILFDLQGLPLNPLFLAHLQMYNGWGNKVRVILAGNAFTA 120	
Qy	121 GKIIISCIPTCFEAAQNTSIAQATMFPVHVIADVRVLEPIEVLDPVLRVLFHNDNAPTMR 180	
Db	121 GKVIICCPVPGFQTLISIAQATLFPHLIADVRLEPIEVLDPVLRVLTHTNDNQPTMR 180	
Qy	181 LVCMLYTPLRASGSSGTDPEFVIAGRVLTCPSPDFSEFLVPPNVEQKTRPFSVPNLPN 240	
Db	181 LLCMLYTPLRGGSGGTDAFVAGRVLTCPSSDFNLFVLPPTVEQKTRPFSVPNIPLO 240	
Qy	241 TLSNSRVPSLIKMMVSRDHQVQFONGRVTLDGQLQGTPTTSASOLCKIRGSVFHANG 300	
Db	241 LLSNSRVPLTQSVLSPDQAQNVQFQNGRGTDTGQLLGTTPVSVSILKFRGKV--SAG 298	
Qy	301 GNGYNLTLDGSPYHAFESPAPIGFDPDLGECDHMEAS-PTQPTNGDVIKQINVKOESA 359	
Db	299 SKVINLTLDGSPFLAFEPAPTGPDLGTDHMEIENSLNSQSSGNPILLRDIQPNSS 358	
Qy	360 -FAPHLGTIQADGLSDSVNTNMIKLGW-----VSPVSDGHRGDVDPWIPRYGSTLT 412	
Db	359 DFVPHLGSVSVTTAIDTAGDT--LGTIQWTSQPSNVTVP-----DVNFWTIPQYGSSLA 411	
Qy	413 EAAQLAPIPPGGEAIVFMSDFP---TAHGTNGLSVPTIPOEEFTHFVNEQAPTRG 469	
Db	412 EASQLAPVVPYPPGGEAIVFMSIPGPNTHAKPN--LVPCLDQEEFTHFVSEQAPSMG 469	
Qy	470 EAALLHYLDPTHNLNLFKLYPEGFMTCVPSNCGTGPQTLPINGVFVFSWSRFRYQLK 529	
Db	470 EAAALVHYVDPTNRLNLFKLYPEGFTICVFN--GTGPOQLPLNGVFVFASWSRFRYQLK 527	
Qy	530 PVGTAGPA-CRLGIRR 544	
Db	528 PVGTASSARGRLGVRR 543	

Search completed: March 10, 2003, 18:27:37
Job time : 40 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2003, 18:18:50 ; Search time 18 Seconds
(without alignments)
1255.811 Million cell updates/sec

Title: US-09-926-799-1
Perfect score: 2896
Sequence: 1 MMASKDAPNTMDGTSGAGQ.....YQLKPVGTAGPACRLGIRRS 545

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1974.5	68.2	546	1 COAT_SOUV3	Q04542 southampton
2	1148.5	39.7	539	1 COAT_LORDV	P54635 lordsdale v
3	316	10.9	2344	1 POLN_RHDDV	P27410 rabbit hemo
4	300.5	10.4	702	1 COAT_SMSV1	P36284 san miquel
5	289	10.0	2208	1 POLN_MANCV	Q69014 manchester
6	276	9.5	671	1 COAT_FCVF9	P27406 feline cali
7	274	9.5	668	1 COAT_FCVF4	P27405 feline cali
8	269	9.3	688	1 COAT_FCVF6	P27404 feline cali
9	264	9.1	703	1 COAT_SMSV4	P36285 san miquel
10	146.5	5.1	2205	1 POLG_POL2W	P23069 poliovirus
11	145.5	5.0	2207	1 POLG_POL2L	P06210 poliovirus
12	145	5.0	2206	1 POLG_POL3L	P03302 poliovirus
13	143.5	5.0	2194	1 POLG_HE701	P32337 human enter
14	139.5	4.8	2206	1 POLG_POL32	P06209 poliovirus
15	137.5	4.7	2209	1 POLG_POL1S	P03301 poliovirus
16	136.5	4.7	2208	1 POLH_POL1M	P03300 poliovirus
17	135.5	4.7	2206	1 POLG_POL1M	P03299 p genome po
18	133.5	4.6	2179	1 POLG_HRV14	P03303 human rhino
19	124.5	4.3	855	1 POLG_HRV3	Q82081 human rhino
20	124	4.3	2214	1 POLG_CXA24	P36290 c genome po
21	123.5	4.3	2175	1 POLG_BOVEV	P12915 bovine ente
22	119.5	4.1	2196	1 POLG_EC05N	Q9Y1J1 e genome po
23	118.5	4.1	2194	1 POLG_EC30B	Q9WJ78 e genome po
24	117.5	4.1	2185	1 POLG_CXB5P	Q03053 c genome po
25	117	4.0	788	1 BCSE_XAN4E	P58933 xanthomonas
26	117	4.0	2183	1 POLG_CXB4E	Q86887 c genome po
27	117	4.0	2183	1 POLG_CXB4J	P08292 c genome po
28	116.5	4.0	2206	1 POLG_CXA21	P22055 c genome po
29	115	4.0	2185	1 POLG_SVDVH	P16604 s genome po
30	115	4.0	2185	1 POLG_SVDVU	P13900 s genome po
31	113.5	3.9	2184	1 POLG_HRV89	P07210 human rhino
32	112.5	3.9	2184	1 POLG_EC01F	Q9Y1734 e genome po
33	112	3.9	2195	1 POLG_EC11G	P29813 e genome po

34	111.5	3.9	516	1 MEPA_XENLA	Q03414 xenopus lae
35	111	3.8	2153	1 POLG_HRV16	Q82122 human rhino
36	109.5	3.8	1022	1 TLD_BRARE	O57460 brachydanio
37	109	3.8	1131	1 YPO4_CAEEL	P53015 caenorhabdi
38	109	3.8	2157	1 POLG_HRV1B	P12916 human rhino
39	107.5	3.7	862	1 POLG_EC16H	Q66790 echovirus 1
40	106	3.7	1493	1 M3K1_MOUSE	P53349 mus musculu
41	106	3.7	2168	1 POLG_PSV9U	O41174 p genome po
42	106	3.7	2332	1 POLG_FMDVO	P03305 f genome po
43	105.5	3.6	2201	1 POLG_CXA9	P21404 c genome po
44	105	3.6	2021	1 OMPA_RICCN	Q52657 rickettsia
45	105	3.6	2193	1 POLG_HE71M	Q66479 human enter

ALIGNMENTS

RESULT 1
COAT_SOUV3
ID COAT_SOUV3 STANDARD; PRT; 546 AA.
AC Q04542;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 01-OCT-1994 (Rel. 30, Last annotation update)
DE Coat protein (Capsid protein).
OS Southampton virus (serotype 3).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=37129;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93142023; PubMed=8380940;
RA Lambden P.R., Caul E.O., Ashley C.R., Clarke I.N.;
RT "Sequence and genome organization of a human small round-structured (Norwalk-like) virus."
RL Science 259:516-519(1993).
CC -!- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L07418; AAA92984.1; -
CC PIR; B37491; B37491.
CC InterPro; IPR004005; Calici_coat.
CC Pfam; PF00915; Calici_coat; 1.
KW Coat protein; Glycoprotein.
FT CARBOHYD 73 73 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 303 303 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 340 340 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 441 441 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 546 AA; 58774 MW; C82B2A85AD4B05EA CRC64;

Query Match 68.2%; Score 1974.5; DB 1; Length 546;
Best Local Similarity 66.6%; Pred. No. 2.1e-143;
Matches 367; Conservative 75; Mismatches 96; Indels 13; Gaps 7;

QY	1	MMASKDAPNTMDGTSGAGOLVPEANTAPISMEPVAGATAATAAGQVNMIDPWTNNY	60
Db	1	MMASKDAQSAGAGAGOLVPEVNTADPLPMEPVAGTPTATAGQVNMIDPWTNNF	60
QY	61	VQAQGEFTISPNNTPCDILFDLQGLPHNLPLSLHQAQNGVGNKVKVLLAGNAFTA	120
Db	61	VQSPQGEFTISPNNTPCDILFDLQGLPHNLPLSLHQAQNGVGNKVKVLLAGNAFTA	120
QY	121	GKIIISCIPPGFAAQNISIAQAQTMFPHVIAVRLVLEPIEVLPEEDVRNVLFNHNDAPTMR	180
Db	121	GKIIICVPPGFTSSSLITIAQAQLFPHVIAVRLVLEPIEVLPEEDVRNVLVHTNDNQPTMR	180

```
QY 181 LVCMLYTLRASGSSGTDPFVIAGRVLTCPSPDFSEFLVPPNVQKTKPFSVPNPLN 240
DB 181 LVCMLYTLRTGGSGNSDSFVWAGRVLTAPSDSFSEFLVPPNIEQKTRAFVNIPLQ 240
QY 241 TLSNRVPSLTKSMVSDHGMVQFQNGRVTLTGQLOGTTPTASOLCKIRGSVFHANG 300
DB 241 TLSNRVPSLTKSMVSDHGMVQFQNGRVTLTGQLOGTTPTASOLCKIRGSVFHANG 300
QY 301 GNGVNLTELDGSPYHAFESPAPGPDGLGECDHMEASPT--TQNTGDIKQINVKOE-S 358
DB 301 GNGVNLTELDGSPYHAFESPAPGPDGLGECDHMEASPT--TQNTGDIKQINVKOE-S 358
QY 359 AFAPHLGTIQADGLSDSVSNVNTNMIKLGWSPVSDGHRGDVDVPIPRYQSTLTERAQLA 418
DB 359 AFAPHLGTIQADGLSDSVSNVNTNMIKLGWSPVSDGHRGDVDVPIPRYQSTLTERAQLA 418
QY 419 PPIYPGGEAIVFMSDFPIAHGTNGLS-----VPCITPOEFVTHFNEQAPTRGEAALL 474
DB 419 PPIYPGGEAIVFMSDFPIAHGTNGLS-----VPCITPOEFVTHFNEQAPTRGEAALL 474
QY 475 HYLDPDTHRNIGERKLYPEGMTCPVNSSGTGPTLPINGVFVSVWSRFYQKLPVGTA 534
DB 475 HYLDPDTHRNIGERKLYPEGMTCPVNSSGTGPTLPINGVFVSVWSRFYQKLPVGTA 534
QY 535 GPA-CRLGIRR 544
DB 535 STARKLGVRR 545
```

```
RESULT 2
COAT_LORDV STANDARD; PRT; 539 AA.
ID COAT_LORDV
AC P54635;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Coat protein (Capsid protein).
OS Lordsdale virus (Human enteric calicivirus).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=82658;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96005060; PubMed=7561776;
RA Dingle K.E., Lambden P.R., Caul E.O., Clarke I.N.;
RT "Human enteric Caliciviridae: the complete genome sequence and
RT expression of virus-like particles from a genetic group II small
RT round structured virus."
RL J. Gen. Virol. 76:2349-2355(1995).
CC -!- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X86557; CAA60255.1; -.
DR InterPro; IPR004005; Calici_coat.
DR Pfam; PF00915; Calici_coat; 1.
KW Coat protein; Glycoprotein.
FT CARBOHYD 172 172 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 307 307 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 406 406 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 531 531 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 539 AA; 58775 MW; 5E5C63E7F2C5FD21 CRC64;
```

Query Match 39.7%; Score 1148.5; DB 1; Length 539;
Best Local Similarity 44.9%; Pred. No. 2.8e-80;
Matches 254; Conservative 77; Mismatches 186; Indels 49; Gaps 16;

```
QY 1 MMWASKADPTNMDGTSAGOLVPEANTAPISMEPVAGAATAAATAGQVNMIDPWIMNNY 60
DB 1 MMWASKADNAPS--DCS--AANLVPEVNN--EVMALEPVVGAIAAPVAGQNVIDPWIRNNE 56
QY 61 VQAPQGEFTISPNTTGDILDLQGLPHLNPFLSHLAQMYNGVGNMVKVYLLAGNAFTA 120
DB 61 VQAPQGEFTVSPNAPGEILWSAPLGPLDNPYLSHLSRMNYAGYGFVQVILAGNAFTA 116
QY 121 GKIIISCIIPGFAAQSTIAQATMFPHVIADVRLBPIEVPLEDVRNVLFF--NNDNAPTM 179
DB 121 GKIIISCIIPGFAAQSTIAQATMFPHVIADVRLBPIEVPLEDVRNVLFF--NNDNAPTM 179
QY 177 KLVIAFAVPNFTTEGLSPSQVTFMFFHIIIVDRQLPEVLIPLDVRNNFYHQAQNDSTL 176
DB 177 KLVIAFAVPNFTTEGLSPSQVTFMFFHIIIVDRQLPEVLIPLDVRNNFYHQAQNDSTL 176
QY 180 RLVCMYLTPLRASGSSGTDPFVIAGRVLTCPSPDFSEFLVPPNVQKTKPFSVPNPL 239
DB 180 RLVCMYLTPLRASGSSGTDPFVIAGRVLTCPSPDFSEFLVPPNVQKTKPFSVPNPL 239
QY 240 NTLNSRVPSLTKSMVSDHGMVQFQNGRVTLTGQLOGTTPTASOLCKIRGSVFHANG 299
DB 240 NTLNSRVPSLTKSMVSDHGMVQFQNGRVTLTGQLOGTTPTASOLCKIRGSVFHANG 299
QY 299 GSHDYTMNLASQNSWVDPTEIPAPLGPDPFGVKIQGLLTQTTRADGSTRAHKAT-VST 353
DB 299 GSHDYTMNLASQNSWVDPTEIPAPLGPDPFGVKIQGLLTQTTRADGSTRAHKAT-VST 353
QY 346 GDVIKQINVKQESAFAPHLGTIQ--ADGLSDSVSNVNTNMIKLGWSPVSDG--HRGDVDP 401
DB 346 GDVIKQINVKQESAFAPHLGTIQ--ADGLSDSVSNVNTNMIKLGWSPVSDG--HRGDVDP 401
QY 402 WTIPTYGSTLTAAQALAPPIYPGGEAIVFMSDFPIAHGTNGLSVPCITPOEFVTHFV 461
DB 402 WTIPTYGSTLTAAQALAPPIYPGGEAIVFMSDFPIAHGTNGLSVPCITPOEFVTHFV 461
QY 462 NEQAPTRGEAALHYLDPDTHRNIGERKLYPEGMTCPVNSSGTGPTQ--TLPINGVFV 519
DB 462 NEQAPTRGEAALHYLDPDTHRNIGERKLYPEGMTCPVNSSGTGPTQ--TLPINGVFV 519
QY 520 SWVSRYFQKLPVGTAGPACRLGIRRS 545
DB 520 SWVSRYFQKLPVGTAGPACRLGIRRS 545
QY 518 SWVNFYTLAPMGNG-----TGRRA 538
DB 518 SWVNFYTLAPMGNG-----TGRRA 538
```

RESULT 3

```
POLN_RHDV STANDARD; PRT; 2344 AA.
ID POLN_RHDV
AC P27410;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Non-structural polyprotein [Contains: RNA-directed RNA polymerase
DE (EC 2.7.7.48); Thiol protease P3C (EC 3.4.22.-); Helicase (2C like
DE protein); Coat protein].
OS Rabbit hemorrhagic disease virus (RHDV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Lagovirus.
OX NCBI_TaxID=11976;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91361557; PubMed=1840711;
RA Meyers G., Wirblich C., Thiel H.-J.;
RT "Rabbit hemorrhagic disease virus -- molecular cloning and nucleotide
RT sequencing of a calicivirus genome."
RL Virology 184:664-676(1991).
CC -!- FUNCTION: THE P3C CYSTEINE PROTEASE IS THE PROTEASE RESPONSIBLE
CC FOR CLEAVAGE AT CERTAIN Q/G SITES IN THE POLYPROTEIN.
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC [RNA](N).
CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C24.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
```

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----
DR EMBL; MG7473; AAA47285.1; -
DR MEROPS; C24.001; -
DR InterPro; IP0004005; Calici_coat.
DR InterPro; IP0004004; Calici_pol_hel.
DR InterPro; IP0003317; Endoptase_C24.
DR InterPro; IP0006005; RNA_helicase.
DR InterPro; IP0001205; RNA_pol_P3D.
DR pfam; PF00680; RNA_dep_RNA_pol; 1.
DR pfam; PF00910; RNA_helicase; 1.
DR pfam; PF00915; Calici_coat; 1.
DR pfam; PF03510; Peptidase_C24; 1.
DR PRINTS; PR00916; ZCENDOPTASE.
DR PRINTS; PR00918; CALICIVIRUS.
KW Polyprotein; Transferase; RNA-directed RNA polymerase; Hydrolase;
KW Thiol protease; Helicase; ATP-binding; Coat protein.
FT CHAIN 7508 7707 HELICASE (P2C-LIKE).
FT CHAIN 1109 1251 PROTEASE P3C.
FT CHAIN ? 2344 RNA-DIRECTED RNA POLYMERASE.
FT CHAIN ? 2344 COAT PROTEIN.
FT NP_BIND 522 529 ATP (POTENTIAL).
FT ACT_SITE 1212 1212 PROTEASE (POTENTIAL).
FT ACT_SITE 1227 1227 PROTEASE (POTENTIAL).
SQ SEQUENCE 2344 AA; 257066 MW; 1454C248F81E9212 CRC64;

Query Match 10.9%; Score 316; DB 1; Length 2344;
Best Local Similarity 28.3%; Pred. No. 8e-16;
Matches 132; Conservative 60; Mismatches 172; Indels 102; Gaps 21;

QY 12 MDGTSAGQLVPEANTAEPTIS-----MEPVAGAAATAA-----ATAG----- 47
DB 1766 MEKRAARAPOGAAGTATTASVPGTTTDDGMDPGVVATTITAE NSSASITAGIGGPPQ 1825
QY 48 QVNMDPWIMNMYVOAPOGETISPNNTPGDILFDLQGLPHLNPFLSHLAQMYNGWVGNM 107
DB 1826 QVDOQETWRTNFYY---NDVFTNSVADAPGSLITYVQHSQNNPFTAVLSQMYAGWAGM 1882
QY 108 KVKVLLAGNAFTAGKIIISCIPGFAAONTISIA---QATMPFHVIADVRLVLEPIEVPLED 164
DB 1883 QFREIVAGSGVFGRLVRAVIPPQ-----IEIGPGLVFRPPHVVIVDARSLEPVTITMPD 1937
QY 165 VRNVLFH-NNDNATMRVLCMLYTPL--RASGSSSGTDPFVIAGRVLTCPSPDFSLFLV 221
DB 1938 LRPNIYHPTGPGVLPTLVLSVNNLNIPFGGSTS-----AIQVTVETRPSEDEFYVIR 1992
QY 222 PPNVEQKTKPFVSPNPLNLTLSNRVPSLTKSMVMVRDGHQMVQFQ-----NGRVT 272
DB 1993 APS--SKTVDSISPAGLITT-----PVLITGVGNDNRWNGQIVGLQVPVGGFSTCNRHN 2044
QY 273 LDGLOQTTPTSASOLKIRGSVFHANGNGYNTLTEL-----DGSPYHAFESP-----APIG 324
DB 2045 LNSTYGVSSPRFGDIDHRRGSASY--SGSNATNVLOFYANAGS---AIDNPISQVAPDG 2100
QY 375 PPDLEGCDWHMEASPTTQNTGDKVQINVKQESAFAPHLCTIOADGLSDVSVNTNIAK 384
DB 2101 PPDMSFVPFNGPGIIPAGWGVFGAIWNSN-----SGAPNVTIVQA-----YE 2142
QY 365 LGWVSPVSDGHRGDVDPWVPIPRYGSTLTEAAQ-LAPPIYPPGFGEA 429
DB 2143 LGFAT---GAPGNLQP-----TTNTSGAQTVAKSIYAVVTGTA 2177

RESULT 4
COAT_SMSV1 STANDARD; PRT: 702 AA.
AC P36284;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Coat protein (Capsid protein).

OS San Miguel sea lion virus (serotype 1) (SMSV 1).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Vesiviruts.
OX NCBI_TaxID=36406;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92410750; PubMed=1529644;
RA Neill J.D.;
RT "Nucleotide sequence of the capsid protein gene of two serotypes of
San Miguel sea lion virus: identification of conserved and non-
conserved amino acid sequences among calicivirus capsid proteins.";
RL Virus Res. 24:211-222(1992).
CC -!- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

CC EMBL; M87481; AAAL6217.1; -
DR PIR; A48562; A48562.
DR InterPro; IP004005; Calici_coat.
DR pfam; PF0915; Calici_coat; 1.
KW Coat protein; Glycoprotein.
FT CARBOHYD 208 208 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 481 481 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 493 493 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 545 545 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 702 AA; 77850 MW; E6E5A58523DEE3D7 CRC64;

Query Match 10.4%; Score 300.5; DB 1; Length 702;
Best Local Similarity 25.2%; Pred. No. 2.6e-15;
Matches 137; Conservative 66; Mismatches 209; Indels 131; Gaps 25;

QY 13 DGTSGAGQLVPEANTAEPTISMEPVAGAAATAAGVNMID-PW-IMNNVVOAPOGEFTI 70
DB 154 DPGGADIVTEEGTGVVQOQPVPAQSALTTLAAASTGKTVDCEWTFSSVHTA---VNW 209
QY 71 SPNNTPGDILFDLQGLPHLNPFLSHLAQMYNGWVGNKVKVLLAGNAFTAGKIIISCIPP 130
DB 210 STTEAOGKILFRLSPNLPYLRHISLSTWSSGGIDVFRFTVSGSVFGGKLAALIVPP 269
QY 131 GFAAONTISIAQATMFHVIADVRLVLEPIEVPLEDVRNVLFHNNNDNAPTMRVLCMLYTPL- 189
DB 270 GI--EPVESPTMLOYPHVLFDARQTEPIFTIPDIRKTLYHSMDDTDTTRELVMVYNELI 327
QY 190 -----RASGSSSGTDPFVIAGRVLTCPSPDFSLFLVPPNVBEOKTKPFSVPN--LPLNT 241
DB 328 NPTEQSEPKSSCSIT-----VETRPSSDFTFSLKPPG--SLLKHGSIPLDIPRNS 377
QY 242 --LJNSRVPLSLIKSMVMVRDGHQMVQFQNGR-----VTLD----- 274
DB 378 RHWGMGNWSTIDGVV-----QPRVFQSNRHFDFDSTTTGWSTPPYIPIEVLLEKLDRG 432
QY 275 GOLQGTTPTSASOL-----CKIRGSVFHANGNGYNTLTEL-----DGSPYHAFESPA 321
DB 433 GQIFKVTDTREKSLVPLGPDGWPDTTITPTAMTASNGNDYTVAEYRITNNGTHFGFYIMG 492
QY 322 PI-----GFPDLGECDDHMEASPTTQNTGDKVQINVKQESAFAP-----HLGTIQA 369
DB 493 NLFTKVKGSDNLGET---QQTSTRLEFASVG-----NYKQNTINPTHKITSNSLVVYDA 543
QY 370 DGLSDSVNVT-----NMTAKLGWV-----SPVSDGHRGDVDPWVPIPRYGSTLTEAAQLA--P 419
DB 544 NNVSAATAKTTHSTHSTMSHLGYLVLDVDESPV-----GSDSTKVYRIATLP 587
QY 420 PIYPGFGEAIVRFMSDFPIAHGTNGLSVPCPTIPOEFVTHFVNQOAPTRCEAAALLHY-LD 478
DB 588 EAFYNG-GNFPVFTFNKIQIGH-----FDRHTKCFNSOVLMTSOKLAENHTLTP 636

```

QY 479 PDT 481
DB 637 PDS 639

RESULT 5
POLN_MANCV STANDARD; PRT: 2208 AA.
AC 069014;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE genome polyprotein [Contains: RNA-directed RNA polymerase
DE (EC 2.7.7.48); Thiol protease 3C (EC 3.4.22.-); Helicase (2C like
DE protein); Coat protein].
OS Manchester virus (human enteric calicivirus).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Sapovirus.
OX NCBI_TaxID=82659;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95390791; PubMed=7661689;
RA Liu B.L., Clarke I.N., Gaul E.O., Lambden P.R.;
RT "Human enteric caliciviruses have a unique genome structure and are
RL distinct from the Norwalk-like viruses.";
Arch. Virol. 140:1345-1356(1995).
CC -!- FUNCTION: P2C IS IMPORTANT IN RNA REPLICATION (BY SIMILARITY).
CC -!- FUNCTION: THE CYSTEINE PROTEASE IS THE PROTEASE RESPONSIBLE FOR
CC THE POST-TRANSLATIONAL PROCESSING OF THE POLYPEPTIDE (BY
CC SIMILARITY).
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC (RNA)(N).
CC -!- PPM: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -!- SIMILARITY: TO PICORNAVIRUS POLYPEPTIDES.
CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C24.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; X86560; CAA60262.1;
CC MEROPS; C24 UPW;
CC InterPro; IPR004005; Calici_coat.
CC InterPro; IPR004004; Calici_pol_hel.
CC InterPro; IPR000317; Endoptase_C24.
CC InterPro; IPR000605; RNA_helicase.
CC InterPro; IPR001205; RNA_pol_P3D.
CC Pfam; PF00680; RNA_dep_RNA_pol; 1.
CC Pfam; PF00910; RNA_helicase; 1.
CC Pfam; PF00915; Calici_coat; 1.
CC Pfam; PF03510; Peptidase_C24; 1.
CC PRINTS; PR00916; 2CENDOPTASE.
CC PRINTS; PR00918; CALICIVIRUSNS.
CC PolyProtein; Transferase; RNA-directed RNA polymerase; Hydrolase;
KW Thiol protease; Helicase; ATP-binding; Coat protein.
FT CHAIN ? ? HELICASE (P2C-LIKE).
FT CHAIN ? ? PROTEASE (P3C).
FT CHAIN ? ? RNA-DIRECTED RNA POLYMERASE.
FT CHAIN ? 2208 COAT PROTEIN.
FT NE_BIND 408 415 ATP (POTENTIAL).
FT ACT_SITE 1097 1097 PROTEASE (BY SIMILARITY).
FT ACT_SITE 1112 1112 PROTEASE (BY SIMILARITY).
SQ SEQUENCE 2208 AA; 242736 MW; 3E299D5BA20E45DC CRC64;

Query Match 10.0%; Score 289; DB 1; Length 2208;
Best Local Similarity 22.9%; Pred. No. 9.4e-14;
Matches 138; Conservative 82; Mismatches 24; Indels 138; Gaps 25;

- QY 14 GTSGAGOLVPEANTAEPISEPVAGAA-----TAAATAGQVNMIDPWIMNMYVQAPQGEFT 69

```

```

Db 1670 GTTG-----PTTSHVVVANEQPNGAQRLELAVATGA-----IQSNVPEAIRNCEFA 1716
QY 70 I-----SPNNTPGDILFDLQGLPHLNPFLSHLAQNYNGVGMKVKVLLAGNAFTAGKI 123
Db 1717 VFTFAWNRMTGTFTLGSISLHPNINPYTSHLSGMWAGGGSFVRLSTSGSGVFAGRI 1776
QY 124 IISCIPIPGFAAQNISIAQATMFPHVIADVRVLEPIEVPLEDVRNVLFHNDNA-PTMRLV 182
Db 1777 IASVIPPG--VDPSSIRDPGVLPFAFVDARITEPVSFMPIDVRAVDYHRMDGAEPYCSLG 1834
QY 183 CMUYTPLRASGSSGTDPFVIAGRVLTCPSPDFSEFLFVPP--NVBQKTKPFSVPNPLN 240
Db 1835 FWYQPLLNPFEFTAVSTCWVS--VETKGGDFDCLLRPPGQXQMGVSPGGLPRLC 1892
QY 241 TLSNSRVPSLIKSMVSRDHGQWQV-FQNGRVTLGQLOGTTPTSASQLCKIRGVSFHAN 299
Db 1893 YSRGNRVGGLVGMILVAEHEKQVNRHNSVTFGWTAPVNPMAA-----EIVTNOAHST 1948
QY 300 GNGY-----NLTEL-DGSPYHAFESPAPI-----GFPDLGECDDHWEASPTTQ 342
Db 1949 SRHAWLSIGAQNKGPLFGIPNHFDPDSCASTVVGAMDTSLGGRPSTGVC-----GPAIS 2002
QY 343 F-NTGDVIKQINVKQESAFAPH-----LGTIQADGLSDSV-----N 378
Db 2003 FONNGDVE--NDTPSVMPATYDPLTSGTGVALTNSINPASILALVRISNNDFTSGFAND 2060
QY 379 TNNIAKLGWSPVSDGH-RGDVDPMVIPRPGYSGTLTEAQAQLA-----PPIYPP 424
Db 2061 KVVVQMSWEMTGINOIRGQVTPMSGNTYFTSTGANTLVLMQERMLSYDGHQAIIYSS 2120
QY 425 GFGEAIVFFMSDPFIAHGTNGLSVFCTIPQEEVTHFVNEQAPTRGEAALLHVLDPDTHRN 484
Db 2121 QLERTAEVFONDI-----VNIPENSMVAF-----NVETNSA 2151
QY 485 LGEEFLYPEGFTVCPNSSGCTGQTLPLINGVVFVWSRFSYQLKPVGTA--GPACRLGI 542
Db 2152 SFQIGIRPDGMY-----TGGSGIVNVPLEPETRF-----QYVGILPLSAALSGPSGNMGR 2202
QY 543 RR 544
Db 2203 AR 2204

RESULT 6
COAT_FCVF9 STANDARD; PRT: 671 AA.
ID COAT_FCVF9
AC P27406;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE Coat protein (Capsid protein).
GN CYP76.
OS Feline calicivirus (strain F9) (FCV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Vesivirus.
OX NCBI_TaxID=11981;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92410623; PubMed=1529544;
RA Carter M.J., Milton I.D., Meanger J., Bennett M., Gaskell R.M.,
RA Turner P.C.;
RT "The complete nucleotide sequence of a feline calicivirus.";
RL Virology 190:443-448(1992).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=92339673; PubMed=1633955;
RA Meanger J., Carter M.J., Gaskell R.M., Turner P.C.;
RT "Cloning and sequence determination of the feline calicivirus strain
RL Biochem. Soc. Trans. 20:26S-26S(1992).
CC -!- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.
CC -----

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; M86379; AAA79327.1; -
 DR EM3L; Z11536; CAA77636.1; -
 DR PIR; B43382; VCVWF9.
 DR PIR; C45538; C45538.
 DR InterPro; IPR004005; Calici_coat.
 DR Pfam; PF00915; Calici_coat; 1.
 KW Coat protein; Glycoprotein.
 FT CARBOHYD 177 177 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 304 304 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 439 439 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 459 459 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 618 618 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 671 AA; 73441 MW; 33BE86D8370D5E5 CRC64;

Query Match 9.5%; Score 276; DB 1; Length 671;
 Best Local Similarity 26.4%; Pred. No. 1.8e-13;
 Matches 102; Conservative 54; Mismatches 128; Indels 102; Gaps 17;

QY 13 DGTSGAGQLVPEANT-AEPISEMPVAGAATAAATAGQVNNMIDPWIMNNYVQAPQGEFTI- 70
 DB 127 DGSITA-----PEQGTWGGVIAEPSAQMSTADMATCKSVDSW-----EAFFSFH 173
 QY 71 -----SPNNTPGDILDLQGLPHLNPFLSHLAQMYNGVGNMKVKVLLAGNAFTAGKIII 125
 DB 174 TSVNWSSETQGLKFLKQSLGPNLPYLEHLAKLYAWMSGIEVRESISGSGVFGKLA 233
 QY 126 SCIPPGF-AAQNTSIAQATMFPHVIAVRVLEPIEVLDPVNVLPNNNDNAPTRLVCM 184
 DB 234 IVPPGDPVQSGTSMQLQ---YPHVLFDAQVPEVIFCLPDLRLTLHLMSDITDTTSLIVM 290
 QY 185 LY-----TPLRASGSSGTDPFVIAGRVLTCPSPDFSLFLVPPNVQEKTKPFVSPNPLN 240
 DB 291 VYNDLINPYANDANSSG---IVT--VETKPGDFKFLHLKPPG-----SMLTHG 335
 QY 241 TLSNRPVSLIKSMVSRHQMVQFNGRVTLDGQLGQTPTTSASQLCKIRGSVFHANG 300
 DB 336 SIPSDLIPKTSLSLWGNRYNSDITDF-----VIRPEVFQANR 372
 QY 301 GNGYN-----LTLDGSPYHA-----PESPA-PTGFPDL---GEC--- 331
 DB 373 HFDNQTAGWSTPRFRPISVTITEONGAKLGIGVATDYIVPGIDGWPDTTIPGELIPA 432
 QY 332 -DWHM-----EASPTQFNTGDVIK 350
 DB 433 GDVAITNGTNDITATGYDTADIK 458

RESULT 7
 COAT_FCVF4
 ID COAT_FCVF4 STANDARD; PRT; 668 AA.
 AC P27405;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 01-JUN-1994 (Rel. 29, Last annotation update)
 DE Coat protein (capsid protein).
 OS Feline calicivirus (strain Japanese F4) (FCV).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
 OC Vesivirus.
 OC NCBI_TaxID=11980;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91306470; PubMed=1853578;
 RA Tohya Y., Taniguchi Y., Takahashi E., Utagawa E., Takeda N.,
 RA Miyamura K., Yamazaki S., Mikami T.;
 RT "Sequence analysis of the 3'-end of feline calicivirus genome.";

RL Virology 183:810-814(1991).
 CC -!- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALCIVIRUSES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; F90357; BAA14371.1; -
 DR PIR; B40481; VCVWF9.
 DR InterPro; IPR004005; Calici_coat.
 DR Pfam; PF00915; Calici_coat; 1.
 KW Coat protein; Glycoprotein.
 FT CARBOHYD 177 177 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 301 301 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 304 304 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 399 399 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 459 459 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 615 615 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 668 AA; 73589 MW; 85BBD8CB85804E503 CRC64;

Query Match 9.5%; Score 274; DB 1; Length 668;
 Best Local Similarity 31.0%; Pred. No. 2.6e-13;
 Matches 74; Conservative 39; Mismatches 82; Indels 44; Gaps 9;

QY 23 PEANTA-EPISMEPVAGAATAAATAGQVNNMIDPWIMNNYVQAPQGEFTI--T---SPNNT 75
 DB 133 PEQGTAVGGVIAEPSAQMSTADMATCKSVDSW-----EAFSEFTSVNWSSTSET 183
 QY 76 PGDILFDLQGLPHLNPFLSHLAQMYNGVGNMKVKVLLAGNAFTAGKIIISCIIPPGF-AA 134
 DB 184 QGKILFKQLGSLGPNLPYLEHLKLYAWMSGIEVRESISGSGVFGKLAIVVPGVDV 243
 QY 135 QNISIAQATMFPHVIAVRVLEPIEVLDPVNVLPNNNDNAPTRLVCMVLY---TPLR 190
 DB 244 QSTSMQLQ---YPHVLFDAQVPEVITFDLRLSTLHVMSDITDTTSLVIMVNDLINPYA 300
 QY 191 AGSSGSGTDPFVIAGRVLTCPSPDFSLFLVPPNVQEKTKPFVSPNPLNTLSNRVPS 249
 DB 301 NDSNSSG---IVT--VETKPGDFKFLHLKPPG-----SVLTGHSIPS 339

RESULT 8
 COAT_FCV66
 ID COAT_FCV66 STANDARD; PRT; 668 AA.
 AC P27404;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Coat protein (capsid protein).
 OS Feline calicivirus (strain CFI/68 FIV) (FCV).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
 OC Vesivirus.
 OC NCBI_TaxID=11979;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RX MEDLINE=91374597; PubMed=1716692;
 RA Neill J.D., Reardon I.M., Heinrikson R.L.;
 RT "Nucleotide sequence and expression of the capsid protein gene of
 RT feline calicivirus".
 RL J. Virol. 65:5440-5447(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Neill J.D.;
 RL Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.
 CC -!- PTM: The N-terminus is blocked.
 CC -!- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALCIVIRUSES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

DR InterPro: IPR000199; Cys-protease-3C.
DR InterPro: IPR003138; Pico_P1A.
DR InterPro: IPR000081; Pico_P2A.
DR InterPro: IPR002527; Pico_P2B.
DR InterPro: IPR000605; RNA_helicase.
DR InterPro: IPR001205; RNA_pol_P3D.
DR InterPro: IPR001676; Rhv.
DR Pfam: PF00073; rhv; 3.
DR Pfam: PF00548; Cys-protease-3C; 1.
DR Pfam: PF00680; RNA_dep_RNA_pol; 1.
DR Pfam: PF00910; RNA_helicase; 1.
DR Pfam: PF00947; Pico_P2A; 1.
DR Pfam: PF01552; Pico_P2B; 1.
DR Pfam: PF02226; Pico_P1A; 1.
DR ProDom: PD001125; Cys-protease-3C; 1.
DR ProDom: PD001274; Pico_P2B; 1.
DR ProDom: PD001306; Pico_P2A; 1.
KW polyprotein; coat protein; Core protein; Transferase;
KW RNA-directed RNA polymerase; Hydrolase; Thiol protease; Myristate.
FT CHAIN 2 69
FT CHAIN 70 340
FT CHAIN 341 578
FT CHAIN 579 879
FT CHAIN 880 1028
FT CHAIN 1029 1125
FT CHAIN 1126 1454
FT CHAIN 1455 1541
FT CHAIN 1542 1563
FT CHAIN 1564 1746
FT CHAIN 1747 2207
FT LIPID 2 2
FT ACT_SITE 1710 1710
FT ACT_SITE 1724 1724
SQ SEQUENCE 2207 AA; 245829 MW; 2B1E2070B7D44F99 CRC64;

Query Match 5.08; Score 145.5; DB 1; Length 2207;
Best Local Similarity 18.8; Pred. No. 0.0087;
Matches 106; Conservative 72; Mismatches 193; Indels 193; Gaps 24;

QY 22 VPEANTAPISMEPVAGATAAATAGQVNMIDPWNMNYVQAPQGETISFN-----NTP 76
DB 365 IPEDVTTPID--IPGEVRNMELAEIDTMIPLNLTNRKNTDMYRVELNDAAHSDTP 421
QY 77 GDILEDLQHLNPLFLSH-----LAQYNGWGNMVKVLLAGNATAGKILICPPG 131
DB 422 ---ILCLSLSPASPLAHTMLGETILNYTHWAGSLKFTFLFCGSMATGKLLVSYAPP 478
QY 132 FAAQNISAQATMPHVIAADVRLIEVPLEDVRLVFNH--NDNAP-----TMR 180
DB 479 AEAPK-SRKEAMLGTHVLDIGLQSSCTMVVPWISNTTYROTINDSFTEGYISMFYQTR 537
QY 181 LVCMLYTLRASGSSGTDPEVIAGRVLTCPSPDFSELF-----LVP 222
DB 538 VVVPLSTPKMD-----ILGFVSAC--NDFSVELLRDTHISOEAMPQGLGLDIE 585
QY 223 PNVEQKTPFSPVNPPLNTLSNR-----VPS-----LIKSMVYSRDHQ 262
DB 586 GVVEGVTRNALTPLTPANLPDPTOSSGPAHSKETPALTAVETGATNPVPSDVTQTRH-- 643
QY 263 MVQ-----FQGRVTLGDLGGTPTS-ASQLCKIRSVF----- 296
DB 644 VIQKTRSESTVESFARGACVAITEVDNDAPTKRASKLFSWKVITYKDTVOLRRKLEFF 703
QY 297 -----HANGNGYNLTLDGSPYHAFESPAPTFPPDICEGDWHEASPT 340
DB 704 TVSRFDEMEFTFVVTSNYTDANNGHALNQYQIMYPGAPIPGKWDY---TWQTSNPS 760
QY 341 TQNTGDKVQKINVKQESAFPHLGTIOA-----DGLSDSVNTNMIAKLGNWSPVSDGH 395
DB 761 VFYTYGAPPARISV-----PYVGIANAYSHFYDGFVK----- 793
QY 396 RGDVDPWVIPRYGSLTEAAQLAPDIYPPGCEALVPMDSFPIAHNGINGLSVPTIQQE 455

DB 794 -----PLAQOASTEGDSL-----YGAA-----SLNDF-----GSLAV----- 820
QY 456 FVTHFVNEQAPTRGEAALLHYLDP 479
DB 821 ---RVVNDHNPTKLTSKIRVYMKP 841
RESULT 12
POLG_POL3L
ID POLG_POL3L STANDARD: PRT: 2206 AA
AC P03302; O84783; O84784; O84785; O84786; O84787; O84788; O84789;
AC O84790; O98592; O98593; O98594;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins
DE P2A TO P2C, P3A; Genome-linked protein VP3; Picornain 3C
DE (EC 3.4.22.28) (Protease 3C) (P3C); RNA-directed RNA polymerase P3D
DE (EC 2.7.7.48)].
OS Poliovirus type 3 (strains P3/Leon/37 and P3/Leon 12A[1]B).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Enterovirus
OX NCBI_TaxID=12088;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=P3/Leon/37;
RX MEDLINE=84170338; PubMed=6324200;
RA Stanway G., Hughes P.J., Mountford R.C., Reeve P., Minor P.D.,
RA Schild G.C., Almond J.W.;
RT "Comparison of the complete nucleotide sequences of the genomes of
RT the neurovirulent poliovirus P3/Leon/37 and its attenuated Sabin
RT vaccine derivative P3/Leon 12a1b.";
RL Proc. Natl. Acad. Sci. U.S.A. 81:1539-1543(1984).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=P3/Leon 12A[1]B;
RX MEDLINE=83299339; PubMed=6310508;
RA Stanway G., Cann A.J., Hauptmann R., Hughes P.J., Clarke L.D.,
RA Mountford R.C., Minor P.D., Schild G.C., Almond J.W.;
RT "The nucleotide sequence of poliovirus type 3 Leon 12 alb: comparison
RT with poliovirus type 1.";
RL Nucleic Acids Res. 11:5629-5643(1983).
RN [3]
RP X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS) OF 1-878.
RX MEDLINE=95120467; PubMed=7820548;
RA Grant R.A., Hiramath C.N., Filman D.J., Syed R., Andries K.,
RA Hogle J.M.;
RT "Structures of poliovirus complexes with anti-viral drugs:
RT implications for viral stability and drug design.";
RL Curr. Biol. 4:784-797(1994).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF 1-878.
RA Hiramath C.N., Grant R.A., Filman D.J., Hogle J.M.;
RT "Binding of the antiviral drug win51711 to the Sabin strain of type-3
RT poliovirus - structural comparison with drug-binding in rhinovirus-
RT 14.";
RL Acta Crystallogr. D 51:473-489(1995).
CC -!- FUNCTION: P3C POLYPEPTIDE IS A PROTEASE THAT CLEAVES AT CERTAIN
CC Q/G SITES IN THE POLYPROTEIN. IT MAY BE A CYSTEINE PROTEASE.
CC -!- CATALYTIC ACTIVITY: Selective cleavage of Gln-I-Gly bond in the
CC poliovirus polyprotein. In other picornavirus reactions Glu may be
CC substituted for Gln, and Ser or Thr for Gly.
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC (RNA)(N).
CC -!- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -!- PTM: SPECIFIC ENZYMOLOGICAL CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -!- MISCELLANEOUS: THE SEQUENCE OF STRAIN SABIN VACCINE P3/LEON/37 IS
CC SHOWN.
CC -!- MISCELLANEOUS: THE STRAIN SABIN VACCINE P3/LEON/37 IS THE
CC PROGENITOR OF THE STRAIN SABIN VACCINE P3/LEON 12A[1]B.
CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.


```

DR InterPro: IPR003138; Pico_P1A.
DR InterPro: IPR000081; Pico_P2A.
DR InterPro: IPR002527; Pico_P2B.
DR InterPro: IPR000605; RNA_helicase.
DR InterPro: IPR001205; RNA_pol_P3D.
DR InterPro: IPR001676; Rnv.
DR Pfam: PF000073; rhv; 3.
DR Pfam: PF00548; Cys-protease-3C; 1.
DR Pfam: PF00680; RNA_dep_RNA_pol; 1.
DR Pfam: PF00910; RNA_helicase; 1.
DR Pfam: PF00947; Pico_P2A; 1.
DR Pfam: PF01552; Pico_P2B; 1.
DR Pfam: PF02226; Pico_P1A; 1.
DR ProDom: PD001125; Cys-protease-3C; 1.
DR ProDom: PD001274; Pico_P2B; 1.
DR ProDom: PD001306; Pico_P2A; 1.
DR Polyprotein; Coat protein; Core protein; Transferase;
KW RNA-directed RNA polymerase; Hydrolase; Thiol protease; Myristate.
FT CHAIN 2 69
FT CHAIN 70 319
FT CHAIN 320 561
FT CHAIN 562 871
FT CHAIN 872 1014
FT CHAIN 1015 1113
FT CHAIN 1114 1443
FT CHAIN 1444 1532
FT CHAIN 1533 1554
FT CHAIN 1555 1737
FT CHAIN 1738 2194
FT LIPID 2
FT ACT_SITE 1701 1701
FT ACT_SITE 1715 1715
SQ SEQUENCE 2194 AA; 244590 MW; 15DBAE96EE06673C CRC64;

Query Match 5.0%; Score 143.5; DB 1; Length 2194;
Best Local Similarity 19.4%; Pred. No. 0.012;
Matches 94; Conservative 57; Mismatches 166; Indels 167; Gaps 17;

QY 70 ISPNTPGDILP-----DIQL-GPHLNPFLSLAQWNGWGNMKVKVLLAGNATAGKII 124
DB 390 ISAQMDQLLNPIDLIQLEPLNTLLGNISRYTHWSGLEMTMFCGSEMTTKGLI 449
QY 125 ISCIPTGAAQNIATAQTAPHVIAVRVLEPLEVPLEDVRNVLFNHNDNAPTRLVCM 184
DB 450 ICYTPGSSPTDRM-QAMLAHVVMDEGLQSSITIIIPWISSGSHYR-----M 496
QY 185 LYTPLRASGSSGTDPPFVIAGRVLTCPSPDFSLFLVPPNVEQTKPFSPNPLPLNTLSN 244
DB 497 FNTDAKAINAVG-----YVTC-----FMQTNL----- 519
QY 245 SRVSLIKSMVSRDHGMQVQFNGRVTLTGLOGLCTPTTSASQLCKIRGSYFHANGNGY 304
DB 520 -----VAPGGAADQCYIVGMV---AAKKDF 541
QY 305 NLTELDGSPYHAFESPAPIGPDLGECDDHMEASPTQFNIGDVIKQINVKQESAFAPHL 364
DB 542 NUKLRMDS-----PDIGQSAILPEQAATTOI--GEIVKTVANTVSEIKRAEL 586
QY 365 GTIQADGLSDSVNTNMIKLGWSPVSDGHRGDVDPWIPRYGSTLTLEAAQLAPPIYPP 424
DB 587 GVIPS-----LNAVETGATSNTEP-----EAAIQRTVINMH 618
QY 425 GRGEAIV-----FPMSPDPIAHNGTNGLSVCTTIPQEVTHFVN--EQAPTRGGAAL 473
DB 619 GTAECLVENFLGRSALVCMRSFEYKNHSTSTS---SIQKNFFIWTNLTRELVLRRKREL 675
QY 474 LHYLPDTHRNLFGEFKLYPEGFMTCVPSNGTGPQTLPINGVFVSVWSRYQLKPVGT 533
DB 676 FTYLFEDT-----EITIVPTLRLFSSNSVSGPLNLTQAMIV-----PTGA 718
QY 534 AGPA 537
DB 719 RKPS 722

```

RESULT 14

```

POLG_POL32 STANDARD; PRT; 2206 AA.
AC PF06209;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins
DE P2A TO P2C; P3A; Genome-linked protein VP6; Picornain 3C
DE (EC 3.4.22.28) (Protease 3C) (P3C); RNA-directed RNA polymerase P3D
DE (EC 2.7.7.48)].
OS Poliovirus type 3 (strain 23127).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Enterovirus.
OX NCBI_TaxID=12087;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87010550; PubMed=3020156;
RA Hughes P.J., Evans D.M.A., Minor P.D., Schild G.C., Almond J.W.,
RA Stanway G.;
RT "The nucleotide sequence of a type 3 poliovirus isolated during a
RT recent outbreak of poliomyelitis in Finland.";
RL J. Gen. Virol. 67:2093-2102(1986).
CC -|- FUNCTION: P3C POLYPEPTIDE IS A PROTEASE THAT CLEAVES AT CERTAIN
CC Q/G SITES IN THE POLYPEPTIDE. IT MAY BE A CYSTEINE PROTEASE.
CC -|- CATALYTIC ACTIVITY: Selective cleavage of Gln-I-Gly bond in the
CC poliovirus polyprotein. In other picornavirus reactions Glu may be
CC substituted for Gln, and Ser or Thr for Gly.
CC -|- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC {RNA}(N).
CC -|- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHERAL UNITS,
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -|- PTM: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -|- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to licensed@isb-sib.ch).
CC
CC EMBL; X04468; CAA28155.1; -
CC PIR; A27245; GNNY27.
CC HSSP; Q84730; LPVC.
CC MEROPS; C03.001; -.
CC MEROPS; C03.020; -.
CC
CC InterPro: IPR000199; Cys-protease-3C.
CC InterPro: IPR003138; Pico_P1A.
CC InterPro: IPR000081; Pico_P2A.
CC InterPro: IPR002527; Pico_P2B.
CC InterPro: IPR000605; RNA_helicase.
CC InterPro: IPR001205; RNA_pol_P3D.
CC InterPro: IPR001676; Rnv.
CC Pfam; PF000073; rhv; 3.
CC Pfam; PF00548; Cys-protease-3C; 1.
CC Pfam; PF00680; RNA_dep_RNA_pol; 1.
CC Pfam; PF00910; RNA_helicase; 1.
CC Pfam; PF00947; Pico_P2A; 1.
CC Pfam; PF01552; Pico_P2B; 1.
CC Pfam; PF02226; Pico_P1A; 1.
CC ProDom; PD001125; Cys-protease-3C; 1.
CC ProDom; PD001274; Pico_P2B; 1.
CC ProDom; PD001306; Pico_P2A; 1.
KW Polyprotein; Coat protein; Core protein; Transferase;
KW RNA-directed RNA polymerase; Hydrolase; Thiol protease; Myristate.
FT CHAIN 2 69
FT CHAIN 70 340
FT CHAIN 341 578
FT CHAIN 578 578

```

```
FT CHAIN 579 878 COAT PROTEIN VP1.
FT CHAIN 879 1027 CORE PROTEIN P2A.
FT CHAIN 1028 1124 CORE PROTEIN P2B.
FT CHAIN 1125 1453 CORE PROTEIN P2C.
FT CHAIN 1454 1540 CORE PROTEIN P3A.
FT CHAIN 1541 1562 CORE PROTEIN P3B.
FT CHAIN 1563 1745 GENOME-LINKED PROTEIN VPG.
FT CHAIN 1746 2206 PICORNAIN 3C.
FT CHAIN 1746 2206 RNA-DIRECTED RNA POLYMERASE P3D.
FT LIPID 2 2 MYRISTATE (BY SIMILARITY).
FT ACT_SITE 1709 1709 PROTEASE (POTENTIAL).
FT ACT_SITE 1723 1723 PROTEASE (POTENTIAL).
SQ SEQUENCE 2206 AA; 245731 MW; F226AD85403C37BA CRC64;

Query Match 4.8%; Score 139.5; DB 1; Length 2206;
Best Local Similarity 20.8%; Pred. No. 0.025;
Matches 96; Conservative 64; Mismatches 182; Indels 119; Gaps 21;

QY 22 VPENTAEPTSMPEVAGATAAATAGOVNMIDPWIMNMYVOAQOGFTI---SPNPTPGD 78
DB 365 IPEDVTPPID---IPGEVKNVMELEIDTMIPLNLENTKRTMDMYRVRLSDSANLSGP 421
QY 79 ILFDLQLGPHLNPFLSH-----LAQMYNGWGNMKVLLAGNAFTAGKIIISCIPPGFA 133
DB 422 ILC-LSUSPAADPRLSHTMLGEVLYNTHWAGSLKFTFLFCGSMMATGKLLVAYAPEG-A 479
QY 134 AQNISIAQATMFPHVIADVRLVLEPIEDVRNVLPH--NNDNAPTMRLVCMLY----- 186
DB 480 QPPTSREKMLGTHVINDLGLSSCTVWVPWISNVYTRQTTDSFTGEGYISMFYQTRIV 539
QY 187 TPLRASGSSGTDDEVIAGRVLTCPSPDFSLFL-----VPPNVE-----OKT 229
DB 540 VPL-----STPKAMDMGLGFVSAC--NDFSRLLRDTHISOAMPOGVDDDLITEVAQNA 591
QY 230 KPESVP---NLP-----LNTLS-----NSRVPS 249
DB 592 LALSCLPQSLPDTKASGPAHSKEVPTLTAVETGATNPLVPSDTVTRIVIOQRSSES 651
QY 250 LIKSM-----MVSRDHGO---MVQ---FQNGRVT-----LDGQLQGTTPTSASOLCK 290
DB 652 TIESFFARGACVAIIEDVNEQPATNVQKLFATWRTIKYKDIVLRKLEFFTYSRFDMEFT 711
QY 291 IRGSVFHANGNGVNLTELDGSPHAFESAPGFPDGLGECDDHMEASPTQRTNGTVGVIK 350
DB 712 FVVTANETNSNGHALNOVQIIMYPGAPTCKSWDDY---TWOTSSNPSIFITYGAAPA 768
QY 351 QINVKQESAPAPHLGTTQIA-----DGLSDVSVNNTNIAKLG 386
DB 769 RISV-----PYVGLANAYSHFDGPAKVPKASDANDQVG 802

RESULT 15
POLG_POLIS STANDARD; PRT: 2209 AA.
AC P03301; Q84881; Q84882; Q84883; Q84884; Q84885; Q84886; Q84887;
AC Q84888; Q84889; Q84890;
DT 21-JUL-1986 (Rel. 01, Created)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins
DE P2A TO P2C; P3A; Genome-linked protein VPG; Picornain 3C
DE (EC 3.4.22.18) (Protease 3C) (P3C); RNA-directed RNA polymerase P3D
DE (EC 2.7.7.48)].
OS Poliovirus type 1 (strain Sabin).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Enterovirus.
OX NCEI_TaxID-12082;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=83299876; PubMed=6310545;
RA Nomoto A., Omata T., Toyoda H., Kuge S., Horie H., Kataoka Y.,
RA Genba Y., Nakano Y., Imura N.;
RT "Complete nucleotide sequence of the attenuated poliovirus Sabin 1
strain genome.";
```

```
RL Proc. Natl. Acad. Sci. U.S.A. 79:5793-5797(1982).
CC -!- FUNCTION: P3C POLYPEPTIDE IS A PROTEASE THAT CLEAVES AT CERTAIN
CC Q/G SITES IN THE POLYPROTEIN. IT MAY BE A CYSTEINE PROTEASE.
CC -!- CATALYTIC ACTIVITY: Selective cleavage of Gln-I-Gly bond in the
CC poliovirus polyprotein. In other picornavirus reactions Glu may be
CC substituted for Gln, and Ser or Thr for Gly.
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate -> N diphosphate +
CC [RNA](N).
CC -!- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHERAL UNITS,
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -!- PTM: SPECIFIC ENZYMIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -!- MISCELLANEOUS: THIS VIRUS IS A LIVE VACCINE STRAIN DERIVED FROM
CC THE MAHONEY STRAIN BY SPONTANEOUS MUTATIONS DURING THE ATTENUATION
CC PROCESS.
CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; V01150; CAA24465.1; -
CC PIR; A03899; GNNY3P.
CC MEROPS; C03.001; -
CC MEROPS; C03.020; -
CC
CC InterPro: IPR000199; Cys-protease-3C.
CC InterPro: IPR003138; Pico_P1A.
CC InterPro: IPR000081; Pico_P2A.
CC InterPro: IPR002527; Pico_P2B.
CC InterPro: IPR000605; RNA_helicase.
CC InterPro: IPR01205; RNA_pol_P3D.
CC InterPro: IPR001676; Rhv.
CC
CC Pfam; PF00073; rhv; 3.
CC Pfam; PF00548; Cys-protease-3C; 1.
CC Pfam; PF00680; RNA_dep_RNA_pol; 1.
CC Pfam; PF00910; RNA_helicase; 1.
CC Pfam; PF00947; Pico_P2A; 1.
CC Pfam; PF01552; Pico_P2B; 1.
CC Pfam; PF02226; Pico_P1A; 1.
CC ProDom; PD001125; Cys-protease-3C; 1.
CC ProDom; PD001274; Pico_P2B; 1.
CC ProDom; PD001306; Pico_P2A; 1.
CC Polyprotein; Coat protein; Core protein; Transferrase;
CC RNA-directed RNA polymerase; Hydrolase; Thiol protease; Myristate.
FT CHAIN 2 69 COAT PROTEIN VP4.
FT CHAIN 70 341 COAT PROTEIN VP3.
FT CHAIN 342 579 COAT PROTEIN VP1.
FT CHAIN 580 881 COAT PROTEIN P2A.
FT CHAIN 882 1030 CORE PROTEIN P2B.
FT CHAIN 1031 1127 CORE PROTEIN P2C.
FT CHAIN 1128 1456 CORE PROTEIN P2A.
FT CHAIN 1457 1543 CORE PROTEIN P3A.
FT CHAIN 1544 1565 GENOME-LINKED PROTEIN VPG.
FT CHAIN 1566 1747 PICORNAIN 3C.
FT CHAIN 1748 2209 RNA-DIRECTED RNA POLYMERASE P3D.
FT LIPID 2 2 MYRISTATE.
FT ACT_SITE 1712 1712 PROTEASE (POTENTIAL).
FT ACT_SITE 1726 1726 PROTEASE (POTENTIAL).
SQ SEQUENCE 2209 AA; 246576 MW; 9EC1EF4D174A28A4 CRC64;

Query Match 4.7%; Score 137.5; DB 1; Length 2209;
Best Local Similarity 19.3%; Pred. No. 0.036;
Matches 108; Conservative 73; Mismatches 193; Indels 187; Gaps 25;

QY 22 VPENTAEPTSMPEVAGATAAATAGOVNMIDPWIMNMYVOAQOGFTTSPNPT--GDI 79
DB 366 LPEFDVTPPID---IPGEVKNMMELEIDTMIPIFDLSAKKNTMEMYRVRLSKPHTDDP 422
QY 80 LFDLQLGPHLNPFLSH-----LAQMYNGWGNMKVLLAGNAFTAGKIIISCIPPGFAA 134
```

Db 423 ILCLSLPASOPRUSHMLGELINYYTHWAGSLKFTFLFCGSMATGKLLVSYAPPG-AD 481
Qy 135 QNISIQAQMEPHVIAOVRLIEVPLEDVNRNVLPHNN-DNAPT-----MRLVC 183
Db 482 PPKRKEAMLGTHVINDIGLOSSCTMVVPLSNTTYRQTIDDSFTEGGYLSVFIQTRIVV 541
Qy 184 MLYTPLRASGSSGTDPEFVIAGRVLTCPSPDFSF-LFLVPPNVEQK----- 228
Db 542 PLSTPREMD-----ILGFVSAC--NDFSVRMLMRDTHIEQKALAQGLQGMLESMI 589
Qy 229 -----TKPFSVPNPLNLTLSNR-VPS-----LIKSMVSRDHGMVQ 265
Db 590 DNTVRETGAATSRDALENTEASGPAHSKEIPALTAVETGATNPLVPSDTVQTRH--VVQ 647
Qy 266 -----FONGR-----VTLDGQLQGTTPTSASOLCKIRGS----- 294
Db 648 HRSRSESSIESFFARGACVAILIIVDNSASTKNKDKLFTVWKIITYKDTVQLRRKLEFTYS 707
Qy 295 -----VFHAN---GGNGYNLTLDGSPYHAFESPAPIGFPDLGECDDHMEASPTQF 343
Db 708 RFDMEFTFVVTANFTETNNGHALNQVYQIMVPPGAPVPEKWDDY---TWQTSSNPSIFY 764
Qy 344 NTGDVIKQINVKQESAFAPHLGTIQA-----DGLSDYSVNTNMIAKLGWVSPVSDGHRGD 398
Db 765 TYGTAPARISV-----PYVGISNAYSIFYDGFSGV-----PLKD----- 798
Qy 399 VDPWVIPRYGSTLTAAQLAPPIYPGFGGEAIVFFMSDFPIAHGNTGLSVPCWIPQEFVT 458
Db 799 -----QSAALGDSLY---GAA---SLNDFGI-----LAV 821
Qy 459 HFVNEQAPTRGEAALLHYLDP 479
Db 822 RVVNDHNPTKVTSKIRVYLKP 842

Search completed: March 10, 2003, 18:26:52
Job time : 26 secs

XX
PT Kit for the detection and typing of small round-structured virus (SRSV)
PT strains for investigation of food poisoning outbreaks, contains
PT antibodies -

```

XX PS Claim 1; Page 40-42; 84pp; Japanese.
XX CC This invention relates to a kit for the detection and typing of small
XX CC round structured virus (SRSV) strains. The kit contains antibodies
XX CC directed against peptides represented in sequences AAB49700 - AAB49710,
XX CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -
XX CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is
XX CC used for detecting and typing strains of SRSV in order to prevent the
XX CC spread of infection and to examine the epidemiology of outbreaks.
XX SQ Sequence 545 AA;

Query Match 100.0%; Score 2896; DB 22; Length 545;
Best Local Similarity 100.0%; Pred. No. 3, le-250;
Matches 545; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MMASKDAPTNMDGTSAGQLVPEANTAEPI SMEPVAGATAAATAGQVNMIDPWIMNNY 60
DB 1 MMASKDAPTNMDGTSAGQLVPEANTAEPI SMEPVAGATAAATAGQVNMIDPWIMNNY 60

QY 61 VOAPQGEFTISPNNTPGDILFDLQGLPHLNPFLSHLAQMYNGWGMKVKVLLAGNAFTA 120
DB 61 VOAPQGEFTISPNNTPGDILFDLQGLPHLNPFLSHLAQMYNGWGMKVKVLLAGNAFTA 120

QY 121 GKIIISCIPPGFAAGNISIAQATMFPHVIADRVLEPIEVPLEDVNRVLFHNNNDNAPTMR 180
DB 121 GKIIISCIPPGFAAGNISIAQATMFPHVIADRVLEPIEVPLEDVNRVLFHNNNDNAPTMR 180

QY 181 LVCMLYTLPLRASGSSGTDPEVIAGRVLTCPSPDFSLFLVPPNVEQTKPFESVPLNPLN 240
DB 181 LVCMLYTLPLRASGSSGTDPEVIAGRVLTCPSPDFSLFLVPPNVEQTKPFESVPLNPLN 240

QY 241 TILSNRVPSELKSMVSRDHQMGVQFONGRVTLTGQGLGTTPTTSASQLCKIRGSVFHANG 300
DB 241 TILSNRVPSELKSMVSRDHQMGVQFONGRVTLTGQGLGTTPTTSASQLCKIRGSVFHANG 300

QY 301 GNGYNLTLDGSPYHAFESPAPIGFDPDGCEDHMEASPTTFNTGVDVTKQINVKQESAF 360
DB 301 GNGYNLTLDGSPYHAFESPAPIGFDPDGCEDHMEASPTTFNTGVDVTKQINVKQESAF 360

QY 361 APHLGTTIQADGLSDSVNTNMIKLGWVSPVSDGHRGDVDPPWIPRYGSTLTEAAQLAPP 420
DB 361 APHLGTTIQADGLSDSVNTNMIKLGWVSPVSDGHRGDVDPPWIPRYGSTLTEAAQLAPP 420

QY 421 IYPPGGEAIVFMSDFPIAHGTCNLSVCTIPQEFVTHFVNEQAPTRGEALLHYLDPD 480
DB 421 IYPPGGEAIVFMSDFPIAHGTCNLSVCTIPQEFVTHFVNEQAPTRGEALLHYLDPD 480

QY 481 THRLNLEKFLYPEGFMTCVPNSSGTGPTLPIINGVFVFSWSRFFYQKLPVGTAGPACRL 540
DB 481 THRLNLEKFLYPEGFMTCVPNSSGTGPTLPIINGVFVFSWSRFFYQKLPVGTAGPACRL 540

QY 541 GTRRS 545
DB 541 GTRRS 545

RESULT 2
AAB49703
ID AAB49703 standard; protein; 544 AA.
XX AC AAB49703;
XX DT 04-APR-2001 (first entry)
XX DE Small round structured virus protein SEQ ID 4.
XX KW Small round structured virus; SRSV; food poisoning.
XX OS Small round structured virus.
XX PN WO200079280-A1.

```

```

XX PD 28-DEC-2000.
XX PF 22-JUN-2000; 2000WO-JP04095.
XX PR 22-JUN-1999; 99JP-0175928.
XX PA (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.
XX PA (DENK-) DENKA SEIKEN KK.
XX PI Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;
XX DR WPI: 2001-080848/09.
XX DR N-PSDB: AAF29144.
XX PT Kit for the detection and typing of small round-structured virus (SRSV)
XX PT strains for investigation of food poisoning outbreaks, contains
XX PT antibodies -
XX PS Claim 1; Page 47-49; 84pp; Japanese.
XX CC This invention relates to a kit for the detection and typing of small
XX CC round structured virus (SRSV) strains. The kit contains antibodies
XX CC directed against peptides represented in sequences AAB49700 - AAB49710,
XX CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -
XX CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is
XX CC used for detecting and typing strains of SRSV in order to prevent the
XX CC spread of infection and to examine the epidemiology of outbreaks.
XX SQ Sequence 544 AA;

Query Match 68.3%; Score 1979; DB 22; Length 544;
Best Local Similarity 66.8%; Pred. No. 3, le-168;
Matches 368; Conservative 68; Mismatches 101; Indels 14; Gaps 6;

QY 1 MMASKDAPTNMDGTSAGQLVPEANTAEPI SMEPVAGATAAATAGQVNMIDPWIMNNY 60
DB 1 MMASKDAPTNMDGTSAGQLVPEANTAEPI SMEPVAGATAAATAGQVNMIDPWIMNNY 60

QY 61 VOAPQGEFTISPNNTPGDILFDLQGLPHLNPFLSHLAQMYNGWGMKVKVLLAGNAFTA 120
DB 61 VOAPQGEFTISPNNTPGDILFDLQGLPHLNPFLSHLAQMYNGWGMKVKVLLAGNAFTA 120

QY 121 GKIIISCIPPGFAAGNISIAQATMFPHVIADRVLEPIEVPLEDVNRVLFHNNNDNAPTMR 180
DB 121 GKIIISCIPPGFAAGNISIAQATMFPHVIADRVLEPIEVPLEDVNRVLFHNNNDNAPTMR 180

QY 181 LVCMLYTLPLRASGSSGTDPEVIAGRVLTCPSPDFSLFLVPPNVEQTKPFESVPLNPLN 240
DB 181 LVCMLYTLPLRASGSSGTDPEVIAGRVLTCPSPDFSLFLVPPNVEQTKPFESVPLNPLN 240

QY 241 TILSNRVPSELKSMVSRDHQMGVQFONGRVTLTGQGLGTTPTTSASQLCKIRGSVFHANG 300
DB 241 TILSNRVPSELKSMVSRDHQMGVQFONGRVTLTGQGLGTTPTTSASQLCKIRGSVFHANG 300

QY 301 GNGYNLTLDGSPYHAFESPAPIGFDPDGCEDHMEASPTTFNTGVDVTKQINVKQES 358
DB 301 GNGYNLTLDGSPYHAFESPAPIGFDPDGCEDHMEASPTTFNTGVDVTKQINVKQES 358

QY 358 QRVNLTLTGSPYHAFESPAPIGFDPDGCEDHMEASPTTFNTGVDVTKQINVKQES 358
DB 358 QRVNLTLTGSPYHAFESPAPIGFDPDGCEDHMEASPTTFNTGVDVTKQINVKQES 358

QY 359 AFAPHLGTTIQADGLSDSVNTNMIKLGWVSPVSDGHRGDVDPPWIPRYGSTLTEAAQLA 418
DB 359 AFAPHLGTTIQADGLSDSVNTNMIKLGWVSPVSDGHRGDVDPPWIPRYGSTLTEAAQLA 418

QY 418 QVPHLSITLTD--ENVSSGGDVI GTTQWTSPPSDSGGANTNFWKIPDYGSGLAEASQLA 416
DB 418 QVPHLSITLTD--ENVSSGGDVI GTTQWTSPPSDSGGANTNFWKIPDYGSGLAEASQLA 416

QY 419 PPIYPPGGEAIVFMSDFPIAHGTCNLSVCTIPQEFVTHFVNEQAPTRGEALL 474
DB 419 PPIYPPGGEAIVFMSDFPIAHGTCNLSVCTIPQEFVTHFVNEQAPTRGEALL 474

QY 474 PAVYPPGGEAIVFMSDFPIAHGTCNLSVCTIPQEFVTHFVNEQAPTRGEALL 473
DB 474 PAVYPPGGEAIVFMSDFPIAHGTCNLSVCTIPQEFVTHFVNEQAPTRGEALL 473

QY 475 HYLDPDTHRLNLEKFLYPEGFMTCVPNSSGTGPTLPIINGVFVFSWSRFFYQKLPVGT 534
DB 475 HYLDPDTHRLNLEKFLYPEGFMTCVPNSSGTGPTLPIINGVFVFSWSRFFYQKLPVGT 534

QY 534 HYVDPDTHRLNLEKFLYPEGFMTCVPNSSGTGPTLPIINGVFVFSWSRFFYQKLPVGT 533
DB 534 HYVDPDTHRLNLEKFLYPEGFMTCVPNSSGTGPTLPIINGVFVFSWSRFFYQKLPVGT 533

QY 535 GPA-CRLGIRR 544
DB 535 GPA-CRLGIRR 544

```


Qy 61 VQAPGGEFTISPNTPGDIILFDLQGLHPLNPFSLHQAQMYNGVGNMKVLLAGNAFTA 120
 Db 61 VQAPGGEFTISPNTPGDIILFDLQGLHPLNPFSLHQAQMYNGVGNMKVLLAGNAFTA 120
 Qy 121 GKIIISCIIPGFAAQNISIAQATMFPHVIADVRVLEPIEVPLEDVNRVLFHNND-NAPTM 179
 Db 121 GKIIISCIIPGFAAQNISIAQATMFPHVIADVRVLEPIEVPLEDVNRVLFHNND-NAPTM 179
 Qy 180 RLVCMLYTLPLRASGSSGTDGDFVIAQVLTCPSPDFSLFLVPPNVEQTKPFSVPLNPL 239
 Db 180 RLVCMLYTLPLRASGSSGTDGDFVIAQVLTCPSPDFSLFLVPPNVEQTKPFSVPLNPL 239
 Qy 240 NTLNSRVPLSLIKSMVSRDGHQVQFQNGRVTLDGLOGTTPTSASOLCKIRGSVFHAN 299
 Db 240 NTLNSRVPLSLIKSMVSRDGHQVQFQNGRVTLDGLOGTTPTSASOLCKIRGSVFHAN 299
 Qy 239 SSLNSRAPLPISSMGISPDNVQSVQFQNGRCTLGRLVGTTPVSLSHVAKIRGT---S 294
 Db 239 SSLNSRAPLPISSMGISPDNVQSVQFQNGRCTLGRLVGTTPVSLSHVAKIRGT---S 294
 Qy 300 GGNVNLTLDELSPYHAFESPAPIGPDPDLGCDWHMEASPTTQFNIGDVIKQINVKQESA 359
 Db 300 GGNVNLTLDELSPYHAFESPAPIGPDPDLGCDWHMEASPTTQFNIGDVIKQINVKQESA 359
 Qy 295 NGTVINLTDELDTGPFHPFEGPAPIGPDPDLGCDWHN---MTQFCHSSQTOYDVTTPDT 351
 Db 295 NGTVINLTDELDTGPFHPFEGPAPIGPDPDLGCDWHN---MTQFCHSSQTOYDVTTPDT 351
 Qy 360 FAPHLGTQADGLSDVSVNTNMIAKLWSPVSDGHRGDVDPWIPRYGSTLTEAAQALAP 419
 Db 360 FAPHLGTQADGLSDVSVNTNMIAKLWSPVSDGHRGDVDPWIPRYGSTLTEAAQALAP 419
 Qy 352 FVPHLGSIQANGIG---SGNVVGLSWISPPSHSPSGQVLDLKKIPNYGSSITEATHLAP 407
 Db 352 FVPHLGSIQANGIG---SGNVVGLSWISPPSHSPSGQVLDLKKIPNYGSSITEATHLAP 407
 Qy 420 PIYPPGFGAIVFMSDFPIAHGTNGLSVPCPTIQEFVTHFVNEQAPTRGEAALLHYLDP 479
 Db 420 PIYPPGFGAIVFMSDFPIAHGTNGLSVPCPTIQEFVTHFVNEQAPTRGEAALLHYLDP 479
 Qy 408 SVYPPGGEVLVFFSKMP---GPGAYNLPCLLPQEVISHLASEQAPTGVGEAALLHYVDP 464
 Db 408 SVYPPGGEVLVFFSKMP---GPGAYNLPCLLPQEVISHLASEQAPTGVGEAALLHYVDP 464
 Qy 480 DTHRNLEGEKLYPGEFMTCPVNSSTGTPOTLPINGVFVSWVSRFYOLKPVGTAGPA-C 538
 Db 480 DTHRNLEGEKLYPGEFMTCPVNSSTGTPOTLPINGVFVSWVSRFYOLKPVGTAGPA-C 538
 Qy 465 DTGRNLGEFKAYPDGFLTCVPGNASSGQQPLPINGVFVSWVSRFYOLKPVGTASSARG 524
 Db 465 DTGRNLGEFKAYPDGFLTCVPGNASSGQQPLPINGVFVSWVSRFYOLKPVGTASSARG 524
 Qy 539 RLGIIR 544
 Db 525 RLGLRR 530

RESULT 5
 AAB49701 standard; protein; 530 AA.
 XX AAB49701;
 XX AAB49701;
 DT 04-APR-2001 (first entry)
 XX Small round structured virus protein SEQ ID 2.
 XX Small round structured virus; SRSV; food poisoning.
 XX Small round structured virus.
 XX WO200079280-A1.
 XX 28-DEC-2000.
 XX 22-JUN-2000; 2000WO-JP04095.
 XX 22-JUN-1999; 99JP-0175928.
 XX (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.
 XX (DENK-) DENKA SEIKEN KK.
 XX Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;
 WPI; 2001-080848/09.
 XX DR N-PSDB; AAF29142.
 XX Kit for the detection and typing of small round-structured virus (SRSV)
 PT strains for investigation of food poisoning outbreaks, contains
 PT antibodies
 XX Claim 1; Page 42-45; 84pp; Japanese.
 XX PS
 XX PN
 XX XX

CC This invention relates to a kit for the detection and typing of small
 CC round structured virus (SRSV) strains. The kit contains antibodies
 CC directed against protein represented in sequences AAB49700 - AAB49710,
 CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -
 CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is
 CC used for detecting and typing strains of SRSV in order to prevent the
 CC spread of infection and to examine the epidemiology of outbreaks.
 XX Sequence 530 AA;
 Qy Match 66.9%; Score 1938; DB 22; Length 530;
 Qy Best Local Similarity 67.0%; Pred. No. 1.6e-164;
 Qy Matches 366; Conservative 63; Mismatches 99; Indels 18; Gaps 7;
 Qy 1 MMASKDAPTNDGCTSGAGOLVPEANTAPISMEPVAGAAATAAGOVNMIDPWIMNNY 60
 Db 1 MMASKDAPTNDGCTSGAGOLVPEANTAPISMEPVAGAAATAAGOVNMIDPWIMNNY 60
 Qy 61 VQAPGGEFTISPNTPGDIILFDLQGLHPLNPFSLHQAQMYNGVGNMKVLLAGNAFTA 120
 Db 61 VQAPGGEFTISPNTPGDIILFDLQGLHPLNPFSLHQAQMYNGVGNMKVLLAGNAFTA 120
 Qy 121 GKIIISCIIPGFAAQNISIAQATMFPHVIADVRVLEPIEVPLEDVNRVLFHNND-NAPTM 179
 Db 121 GKIIISCIIPGFAAQNISIAQATMFPHVIADVRVLEPIEVPLEDVNRVLFHNND-NAPTM 179
 Qy 180 RLVCMLYTLPLRASGSSGTDGDFVIAQVLTCPSPDFSLFLVPPNVEQTKPFSVPLNPL 239
 Db 180 RLVCMLYTLPLRASGSSGTDGDFVIAQVLTCPSPDFSLFLVPPNVEQTKPFSVPLNPL 239
 Qy 240 NTLNSRVPLSLIKSMVSRDGHQVQFQNGRVTLDGLOGTTPTSASOLCKIRGSVFHAN 299
 Db 240 NTLNSRVPLSLIKSMVSRDGHQVQFQNGRVTLDGLOGTTPTSASOLCKIRGSVFHAN 299
 Qy 239 SSLNSRAPLPISSMGISPDNVQSVQFQNGRCTLGRLVGTTPVSLSHVAKIRGT---S 294
 Db 239 SSLNSRAPLPISSMGISPDNVQSVQFQNGRCTLGRLVGTTPVSLSHVAKIRGT---S 294
 Qy 300 GGNVNLTLDELSPYHAFESPAPIGPDPDLGCDWHMEASPTTQFNIGDVIKQINVKQESA 359
 Db 300 GGNVNLTLDELSPYHAFESPAPIGPDPDLGCDWHMEASPTTQFNIGDVIKQINVKQESA 359
 Qy 295 NGTVINLTDELDTGPFHPFEGPAPIGPDPDLGCDWHN---MTQFCHSSQTOYDVTTPDT 351
 Db 295 NGTVINLTDELDTGPFHPFEGPAPIGPDPDLGCDWHN---MTQFCHSSQTOYDVTTPDT 351
 Qy 360 FAPHLGTQADGLSDVSVNTNMIAKLWSPVSDGHRGDVDPWIPRYGSTLTEAAQALAP 419
 Db 360 FAPHLGTQADGLSDVSVNTNMIAKLWSPVSDGHRGDVDPWIPRYGSTLTEAAQALAP 419
 Qy 352 FVPHLGSIQANGIG---SGNVVGLSWISPPSHSPSGQVLDLKKIPNYGSSITEATHLAP 407
 Db 352 FVPHLGSIQANGIG---SGNVVGLSWISPPSHSPSGQVLDLKKIPNYGSSITEATHLAP 407
 Qy 420 PIYPPGFGAIVFMSDFPIAHGTNGLSVPCPTIQEFVTHFVNEQAPTRGEAALLHYLDP 479
 Db 420 PIYPPGFGAIVFMSDFPIAHGTNGLSVPCPTIQEFVTHFVNEQAPTRGEAALLHYLDP 479
 Qy 408 SVYPPGGEVLVFFSKMP---GPGAYNLPCLLPQEVISHLASEQAPTGVGEAALLHYVDP 464
 Db 408 SVYPPGGEVLVFFSKMP---GPGAYNLPCLLPQEVISHLASEQAPTGVGEAALLHYVDP 464
 Qy 480 DTHRNLEGEKLYPGEFMTCPVNSSTGTPOTLPINGVFVSWVSRFYOLKPVGTAGPA-C 538
 Db 480 DTHRNLEGEKLYPGEFMTCPVNSSTGTPOTLPINGVFVSWVSRFYOLKPVGTAGPA-C 538
 Qy 465 DTGRNLGEFKAYPDGFLTCVPGNASSGQQPLPINGVFVSWVSRFYOLKPVGTASSARG 524
 Db 465 DTGRNLGEFKAYPDGFLTCVPGNASSGQQPLPINGVFVSWVSRFYOLKPVGTASSARG 524
 Qy 539 RLGIIR 544
 Db 525 RLGLRR 530
 RESULT 6
 AAR57091
 ID AAR57091 standard; Protein; 530 AA.
 XX AAR57091;
 XX AAR57091;
 DT 05-OCT-1994 (first entry)
 XX Small round virus SRSV/KY/89 capsid protein.
 XX pathogen; acute gastroenteritis; food poisoning;
 XX seafood contamination; diagnostic assay; human calicivirus;
 XX small round virus; SRSV; KY89; Norwalk virus; capsid protein.
 XX Small round virus SRSV/KY/89.
 XX WO9405700-A.


```

PD      28-DEC-2000.
XX
XX      22-JUN-2000; 2000WO-JP04095.
XX
XX      22-JUN-1999; 99JP-0175928.
XX
XX      (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.
XX      (DENK-) DENKA SEIKEN KK.
XX
XX      Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;
XX      WPI; 2001-080848/09.
XX      N-PSDB; AAF29149.
XX
XX      Kit for the detection and typing of small round-structured virus (SRSV)
XX      strains for investigation of food poisoning outbreaks, contains
XX      antibodies.
XX
XX      Claim 1; Page 59-61; 84pp; Japanese.
XX
XX      This invention relates to a kit for the detection and typing of small
XX      round structured virus (SRSV) strains. The kit contains antibodies
XX      directed against peptides represented in sequences AAB49700 - AAB49710,
XX      which are each SRSV strain specific. Polynucleotide sequences AAF20141 -
XX      AAF20151 represent cDNA encoding the strain specific proteins. The kit is
XX      used for detecting and typing strains of SRSV in order to prevent the
XX      spread of infection and to examine the epidemiology of outbreaks.
XX
XX      Sequence 542 AA;
XX
XX      Query Match      40.5%; Score 1172.5; DB 22; Length 542;
XX      Best Local Similarity 46.4%; Pred. No. 5.3e-96;
XX      Matches 257; Conservative 85; Mismatches 185; Indels 27; Gaps 15;

QY      1 MMASKD-APTNDGTSAGQVLPEANTAEPISEMPVAGATAAATAGQVNMIDPWIMNN 59
DB      1 MKMASDAAPSN---DGAASLVPE-GINETMPELVAGASIAAPAGOTNIDPWIRTN 55

QY      60 YVCAPOGETTISPNTTGDILFDLQGLPHLPFLSHLAQMYNGWKNKVKVLLAGNAFT 119
DB      56 FVQANGCEFTVSPRNSGGEILLNLEGLPOLNPYLAHLSRMVYAGGVQVLLAGNAFT 115

QY      120 AGHIIISCIPTPGFAAGNISIAQATMPPHVIADRVLEPIEVPLEDVRLVLFH-NNDNAPT 178
DB      116 AGHILFAATPPNPLVDMTISPAQITMLPHLIVDVRTLEPIMTLPDVRNVFHFNNQOPR 175

QY      179 MRLVCMLYTPLRASGSSGTDPEVIAGRVLTCPSPDFSLFLVPPNVBQTKPSPVNLPL 238
DB      176 MRLVAMLYTPLRSNG--SGDDVFTVSCRVLTRPTDFEFLVPPSVESKTKPFTLILT 233

QY      239 LNVLSRVPSPSLIKSMVSRDHQWQVONGRVTLTGQLOGTTTTSASOLCKIRGSVEHA 298
DB      234 ISBLTNSRFPPIEQLYTAPNETNVVQONGRCTLDGLOGTTTLLSASVCFLOGRTVAD 293

QY      299 NGGN-GYNLTTEL---DGSYPH-AFESPAPIGPDLGECDMHMEASPTTQFTGTV--IKQ 351
DB      294 NGDNWONLLQLYTPNGASYDPTDEVPALGTQDFSGMLYGLVLTQDNVNVSTGEAKNAG 353

QY      352 INYKQESA-FAPHLGTIQADGLSDVS--VNTNMIKLGWSPVSDGHRGDVDPWVIRYG 408
DB      354 IYISTTSKGFTPKIGSI---GLHSITEHVPNQSRFTPVGVAVD-ENTPFPQOVLPHYA 409

QY      409 STCTEAAQLAPPIYPPGFEAIVFEMSDPIAHGTNGLS--VPCTIPQEFYTHFVNSOAP 466
DB      410 GSIALNTNLAPAVAPTFPGEQLLFFRSRVPVCGVQGGQDADFIDCLLPQEWNVHYEQEAP 469

QY      467 TRGEAALLHYLDPDTHRNIGEFKLYPEGPMTCVPNSSGTGTLPINGVFVFSWSRFF 526
DB      470 SQADVALIRVNPDTGRTLFKALHRSRGTIV--SHTGAYPLVVPVPPNGHFRFDSWVNOFY 527

QY      527 QLKPVGTAGPACRL 540
DB      528 SLAPMGITGNGRRRI 541

Query Match      40.5%; Score 1172.5; DB 22; Length 542;
Best Local Similarity 46.4%; Pred. No. 5.3e-96;
Matches 257; Conservative 85; Mismatches 185; Indels 27; Gaps 15;

QY      1 MMASKD-APTNDGTSAGQVLPEANTAEPISEMPVAGATAAATAGQVNMIDPWIMNNY 60
DB      1 MKMASDAAPSNDAAG---LVPEANN-ETMALEPVAGASIAAPLTGQNNIIDPWIRLNF 56

QY      61 VOAFQGETTISPNTTGDILFDLQGLPHLPFLSHLAQMYNGWKNKVKVLLAGNAFTA 120
DB      57 VOAPNGCEFTVSPRNSGGEILLNLEGLPELNPYLAHLSRMVYAGGVQVLLAGNAFTA 116

QY      121 GKIIISCIPTPGFAAGNISIAQATMPPHVIADRVLEPIEVPLEDVRLVLFH-NNDNAPT 179
DB      117 GRLVFAAVPPHPLPNIISPGQITMFPHVIDVRLTLEPVLPLPVDVNRNFFHYNQNEPRM 176

QY      180 RLVCMLKXPLRASGSSGTDPEVIAGRVLTCPSPDFSLFLVPPNVBQTKPSPVNLPL 239
DB      177 RLVAMLYTPLRSNG--SGDDVFTVSCRVLTRPSPDFDFNVLVPPTLESKTKPFTLILT 234

QY      240 NTLNSRVPSPSLIKSMVSRDHQWQVONGRVTLTGQLOGTTTTSASOLCKIRGSVEHAN 299
DB      235 GELTNSRVPSPIDELXTSPNESLVVQONGRCALDGLGLOGTTQLLPTALCSRGRINOKY 294

QY      300 GNGY----NLTELDGSPYH-AFESPAPIGFPD-----LGECDHMEASPTTQFTG 346
DB      295 SCENHVMNQVTNIGTPFDGTGDVPAPLCTPDSFKGLFGLVSQD-HDNAC-----RSH 348

QY      347 DVIRQINVKQESAFAPHLGTIQADGLSDVSNTNMIKLGWSPVSDGHRGDVDPWVIRP 406
DB      349 DAVIATN---SAKTPKLGAIQIGTWEEEDVHINQTKP---TPVGLFENEGFNQWTLN 402

QY      407 XGSLTEAAQLAPPIYPPGFEAIVFEMSDPIAHGTNGLSVPTCTIPQEFYTHFVNEQAP 466
DB      403 YSGALTNNGLAPPVAPTFPGEQILEFRSHIPKGGVADPVIDCLLPQEWIQLHYESAP 462

QY      467 TRGEAALLHYLDPDTHRNIGEFKLYPEGMTCPVNSSGTGTLPINGVFVFSWSRFF 526
DB      463 SOSDVALIRVNPDTGRTLVFLFAKLHRSRGTIVA--NIGSRPIVVPANGIFREDTWNQFY 520

QY      527 QLKPVGTAGPACRLGIRR 544
DB      521 SLAPMGITGN-----GRRR 533

RESULT 10
AAB49708
ID      AAB49708 standard; protein: 542 AA.
XX
XX      AAB49708;
XX
XX      04-APR-2001 (first entry)
XX
XX      Small round structured virus protein SEQ ID 9.
XX
XX      Small round structured virus; SRSV; food poisoning.
XX
XX      Small round structured virus.
XX
XX      WO200079280-A1.
XX

```

RESULT 11
 AAB49704 standard; protein; 539 AA.
 ID AAB49704 standard; protein; 539 AA.
 XX AC AAB49704;
 XX DT 04-APR-2001 (first entry)
 XX DE Small round structured virus protein SEQ ID 5.
 XX KW Small round structured virus; SRSV; food poisoning.
 XX OS Small round structured virus.
 XX PN WO200079280-A1.
 XX PD 28-DEC-2000.
 XX PF 22-JUN-2000; 2000WO-JP04095.
 XX PR 22-JUN-1999; 99JP-0175928.
 XX PA (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.
 XX PA (DENK-) DENKA SEIKEN KK.
 XX PI Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;
 XX PI WPI; 2001-080848/09.
 XX DR N-PSDB; AAF29145.
 XX PT Kit for the detection and typing of small round-structured virus (SRSV)
 XX PT strains for investigation of food poisoning outbreaks, contains
 XX PT antibodies -
 XX PS Claim 1; Page 50-52; 84pp; Japanese.
 XX CC This invention relates to a kit for the detection and typing of small
 CC round structured virus (SRSV) strains. The kit contains antibodies
 CC directed against peptides represented in sequences AAB49700 - AAB49710,
 CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -
 CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is
 CC used for detecting and typing strains of SRSV in order to prevent the
 CC spread of infection and to examine the epidemiology of outbreaks.
 XX SQ Sequence 539 AA;
 Query Match 40.0%; Score 1157.5; DB 22; Length 539;
 Best Local Similarity 45.6%; Pred. No. 1.1e-94;
 Matches 253; Conservative 83; Mismatches 192; Indels 27; Gaps 14;
 QY 1 MMASKDAPTNMDGTSGAGQLVPEANTAEPISEMEPVAGAAATAAGQVNMIDPWIMNY 60
 Db / 1 MMASKDAPTNMDGTSGAGQLVPEANTAEPISEMEPVAGAAATAAGQVNMIDPWIMNY 60
 QY 61 VQAPGCEFTISPNPTGDIILFDLQGLPHLPFLSHLAQMYNGWGMKVKLLAGNAFTA 120
 Db 57 VQAPGCEFTISPNPTGDIILFDLQGLPHLPFLSHLAQMYNGWGMKVKLLAGNAFTA 116
 QY 121 GKIIISCIIPGCAONISIAQATMFHVIADVRLEPIEVLPLEDVRNVLFIH-NNDNAPTM 179
 Db 117 GKIIISCIIPGCAONISIAQATMFHVIADVRLEPIEVLPLEDVRNVLFIH-NNDNAPTM 176
 QY 180 RLVCMYLYTLRASSSGSDTDFVIAGRLVCPSPDFSLFLVPPNVQKTRKPSVPLNPL 239
 Db 177 KLIAMLYTLPLRA--NNAGDDVFTVSCVLRPSPDFDFLVPVTVESRTPFTVPLTV 234
 QY 240 NTLNSRVPISLKMSWSDRGOMVQFQNGRVTLTGLOGTTPTSASOLCKTRGSRVFAN 299
 Db 235 EEMSNRRPIPLEKLYLCPSSAFVQVQNGRCRTDGLVLTQLSALNLCIFRGDYTHIA 294
 QY 300 GNGY--NLTELDSPPYAFES-PAPIGFPD-LGECWHEASPTTQNTGDTVTKQINVK 355

Db 295 GSHDYTMNLASQNSNDYDTEIPAPLGTDFVGKIQGL--TOTREDGSTRAHKATVS 352
 QY 356 QESA-FAPHLGTIOADGLSDVSVNTNMIKLGWSPVSDG--HRGDVDPWIPRYGSTLT 412
 Db 353 TGSVHFTPKLGSVOYTTDTNDFOTGONTKFTPVGVIOGNHNEPOOWVLPNYSGRGT 412
 QY 413 EAAQLAPPITYPPGGEAIVFFMSDFPIAHGTNGLSVPCPIPOEFVTHFVNBOAPTRGEAA 472
 Db 413 HNVHLAPAVAPTEPGEOLLFFRSTMPGSCGYPMNMLDCLLPQEWVQHFCEAAQAQSDVA 472
 QY 473 LLHYLDPTHNLGEPKLYPGEWTCVPNSSGTGPOTL--PINGVVFVSVWSRYQLKP 530
 Db 473 LLRFVNDPTGVLFECKLHKSGYTV----AHTGPHDLVIPPNGYFRDSDWVNOFYTLAP 528
 QY 531 VGTAGPACRLGIRRS 545
 Db 529 MNGA-----GRREA 538
 RESULT 12
 AAU91272 standard; Protein; 548 AA.
 XX AC AAU91272;
 XX DT 18-JUN-2002 (first entry)
 XX DE Norwalk virus associated polynucleotide #1.
 XX KW Norwalk virus; monoclonal antibody; geno group I; geno group II;
 XX KW immunological detection; food; viral infection.
 XX OS Norwalk virus.
 XX PN JP2002020399-A.
 XX PD 23-JAN-2002.
 XX PF 10-JUL-2000; 2000JP-0208151.
 XX PR 10-JUL-2000; 2000JP-0208151.
 XX PA (OSAP) OSAPA PREFECTURE.
 XX PA (NISE) NIPPON SEIBUTSU KAGAKU CENT KK.
 XX PA (IATR) IATR LAB INC.
 XX DR WPI; 2002-287412/33.
 XX PT A monoclonal antibody useful in the immunological detection and
 XX PT diagnosis of Norwalk virus infection -
 XX PS Disclosure; Page 12-13; 24pp; Japanese.
 XX CC The invention describes a monoclonal antibody recognising Norwalk virus,
 CC a capsid protein of Norwalk virus, or a common antigen epitope on the
 CC capsid protein molecule of geno group I and geno group II. The antibody
 CC is useful for immunological detection and quantitative analysis of
 CC Norwalk virus in foods and the serum of infected patients. This sequence
 CC represents a Norwalk virus associated protein described in the invention.
 XX SQ Sequence 548 AA;
 Query Match 40.0%; Score 1157; DB 23; Length 548;
 Best Local Similarity 45.7%; Pred. No. 1.3e-94;
 Matches 254; Conservative 83; Mismatches 177; Indels 42; Gaps 15;
 QY 2 MMASKDAPTNMDGTSGAGQLVPEANTAEPISEMEPVAGAAATAAGQVNMIDPWIMNY 61
 Db 3 MASSRAAFSN-DGAAG--LVPEINN-EAALDPVAGAAIAAPLTGQNIIDPWIMNMFV 57
 QY 62 QAPQCEFTISPNPTGDIILFDLQGLPHLPFLSHLAQMYNGWGMKVKLLAGNAFTA 121

Db 58 QAPGGEFTVSPRNSGCVLLNLNLGPEINPYLAHLARMYNGYAGGFEVQVVLGACNAFTAG 117
 QY 122 KILISCIIPGFAAONISIAQATMFPHVIADVRVLEPIEVPLEDVNRVLFHNDNAPT-MR 180
 Db 118 KILFAAIPPNFIDNLSAAQITMCPHVIDVRVLEPIEVPLEDVNRVLFHNDNAPT-MR 177
 QY 181 LVCMLYTPLRASGSSGCTDPFVIAGRVLTCPSPDFSEFLVPPNVEQKTKPFSVPNLPN 240
 Db 178 LIAMLYTPLRASGSSGCTDPFVIAGRVLTCPSPDFSEFLVPPNVEQKTKPFSVPNLPN 235
 QY 241 TLSNSRVPSLTKSMVSRDGMVQFNGRVLTDLGLOGTTPTSASOLCKIRGSVEHANG 300
 Db 236 EMSNSRPVPLESLHTSPENIVVQCONGRVLTDLGELMGTQQLPSRICAFRGVLTST 295
 QY 301 G-----NGY---NLTELDGSPYHAFES-PAPIGPDL-GECDHMEASPTTQ 342
 Db 286 RASDAQADTRPRFNYYMHVQOLNLTGTPYPAEDIPGLGTPDRGKV-----FGVASQ 350
 QY 343 FNTGDVVIKQINVKQESA---FAPHLGTIQADGLSDVSVNTNMIKLGWVSPVSDG--HRG 397
 Db 351 RNPDSITTRAHEAKVDTTGTAGRTFKLGSLEISTESS-DEPQNPQTRF---TPVGIGVDNEA 406
 QY 398 DVDPRVRYGCTLTAAQAAPPYPPGFGAEIVFPFMSDFPIAHGTNGLSVPTIPQEFV 457
 Db 407 DFQMSLUPDYGQPTTHMNLAPAVAPNFPGEOLLFERSOLSSGGRSNGVLDLPQEW 466
 QY 458 THEVNEQAPTREGAALHYLDPDTHRLNGELFKLYPEGFTCVPNSSGTCPTLPIGVFV 517
 Db 467 QHFEQESAPACTQVALRVYVNDTGRVLFELAKHLKLGFTIAKN--GDSPTVPPNGYER 524
 QY 518 FVSWSRFVYQKPVGT 533
 Db 525 FESWVNPFTYIAPMGT 540

RESULT 13
 AAB49709
 ID AAB49709 standard; protein; 550 AA.
 XX
 AC AAB49709;
 XX
 DT 04-APR-2001 (first entry)
 DE Small round structured virus protein SEQ ID 10.
 XX
 KW Small round structured virus; SRSV; food poisoning.
 XX
 OS Small round structured virus.
 XX
 PN WO200079280-A1.
 XX
 PD 28-DEC-2000.
 PF 22-JUN-2000; 2000WO-JP04095.
 PR 22-JUN-1999; 99JP-0175928.
 XX
 PA (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.
 PA (DENK-) DENKA SEIKEN KK.
 XX
 PI Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;
 XX
 DR WPI; 2001-080848/09.
 DR N-PSDB; AAF29150.
 XX
 PT Kit for the detection and typing of small round-structured virus (SRSV)
 PT strains for investigation of food poisoning outbreaks, contains
 PT antibodies
 XX
 PS Claim 1; Page 62-64; 84pp; Japanese.
 XX
 CC This invention relates to a kit for the detection and typing of small
 CC round structured virus (SRSV) strains. The kit contains antibodies

CC directed against peptides represented in sequences AAB49700 - AAB49710, which are each SRSV strain specific. Polynucleotide sequences AAB20141 - AAB20151 represent cDNA encoding the strain specific proteins. The kit is used for detecting and typing strains of SRSV in order to prevent the spread of infection and to examine the epidemiology of outbreaks.

XX
 SQ Sequence 550 AA;
 Query Match 39.9%; Score 1154.5; DB 22; Length 550;
 Best Local Similarity 45.6%; Pred. No. 2.2e-94;
 Matches 259; Conservative 73; Mismatches 175; Indels 61; Gaps 14;

QY 1 MMMAKSKO-APINMDGTSAGQOLVPEANTAPFISMEPVAGATAAATAAGVNMIDPMTNN 59
 Db 1 MKMASNDRAPSN-----DGAANLVPEAND-EVMALEPVVGASTAAPVQGNLIIDPWIREN 55
 QY 60 VVQAPQGEFTVSPRNSGCVLLNLNLGPEINPYLAHLARMYNGYAGGFEVQVVLGACNAFT 119
 Db 56 FVQAPQGEFTVSPRNSGCVLLNLNLGPEINPYLAHLARMYNGYAGGFEVQVVLGACNAFT 115
 QY 120 AGKTIISCIIPGFAAONISIAQATMFPHVIADVRVLEPIEVPLEDVNRVLFHNDNAPT 178
 Db 116 AGKILFAAAPPNPFIDNLSAAQITMCPHVIDVRVLEPIEVPLEDVNRVLFHNDNAPT 175
 QY 179 MRLVCMYTPLRASGSSGCTDPFVIAGRVLTCPSPDFSEFLVPPNVEQKTKPFSVPNLP 238
 Db 176 MRLVAMLYTPLRASGSSGCTDPFVIAGRVLTCPSPDFSEFLVPPNVEQKTKPFSVPNLP 232
 QY 239 LNTLSNSRVPSLTKSMVSRDGMVQFNGRVLTDLGLOGTTPTSASOLCKIRGSVEHANG 297
 Db 233 LGELNSRFPFAADIMLYTDPNIESIVVQCONGRVLTDLGLOGTTPTSASOLCKIRGSVEHANG 292
 QY 298 -----ANGNGYNLTDELGSPYHAFES-PAPIGPDLGECDHMEASPTTQ 342
 Db 293 TARAADSTDSQARARNHPLHVQVKNLDGTQYDPTDIPAVLGAIQF-----KGTV 342
 QY 343 FNTGDVVIKQINVKQESA---FAPHLGTIQADGLSDVSVNTNMIKLGWVSPVSDG--HRG 385
 Db 343 FGVASQ-RDVSQEQEQGHYATRAHEAHIDTDPKYAPKLGTLIKSGSD-DEFTNQPIRF 400
 QY 386 GWSPVSDGHRGDVDPVPIRYGSTLTAAQAAPPYPPGFGAEIVFPFMSDFPIAHGTNG 445
 Db 401 ---TPVGMDG-NNRWQLPFDYSGRLTLNMLAPAVSPSPGERILFFRSIVPSAGGYS 456
 QY 446 LSVPTIPQEEVTHEVNEQAPTREGAALHYLDPDTHRLNGELFKLYPEGFTCVPNSSGTC 505
 Db 457 GVIDCLIPQEWVQHFOEAPASQSAVALRVYVNDTGRVLFELAKHLKLGFTIAKN--NCGN 514
 QY 506 GPCTLPIGVFVYQKPVGT 533
 Db 515 NPLVVPNGYERFPAWGNQFTYIAPMGT 542

RESULT 14
 AAB49710
 ID AAB49710 standard; protein; 541 AA.
 XX
 AC AAB49710;
 XX
 DT 04-APR-2001 (first entry)
 DE Small round structured virus protein SEQ ID 11.
 XX
 KW Small round structured virus; SRSV; food poisoning.
 XX
 OS Small round structured virus.
 XX
 PN WO200079280-A1.
 XX
 PD 28-DEC-2000.
 PF 22-JUN-2000; 2000WO-JP04095.
 PR
 XX


```

Query Match      11.1%; Score 321; DB 18; Length 579;
Best Local Similarity 28.3%; Pred No. 9, 2e-20;
Matches 132; Conservative 61; Mismatches 171; Indels 102; Gaps 21;

QY 12 MDGTSAGQLVPENATPAEIS-----MEPVAGAAATA-----ATAG----- 47
DB 1 MEGKARAPOGEAAGTATTASVGTGTTDGMDFGVVATTSVITAENSSASIAIATAGIGGPPQ 60
QY 48 QVNMIDPWIMNNYVQAPQGEFTISPNNTPCDILFDLQGLPHLNPFLSHLAQMYNGWYGNM 107
DB 61 QVDQQTWRTNFTY---NDVFTSVADAPGSILYTVQHSPPQNNPFTAVLSQMYAGWAGGM 117
QY 108 KVKVLLAGNAFTACKIIISCIPGFAAQNTISA---QATMFPHVIAADVRLPEIEVPLED 164
DB 118 QRFIVAGSGVFGGRVLAIVIPG-----IEIGPLEVRFQFPHVVIDARSLEPVITIMPD 172
QY 165 VRNVLFH-NNDNAPFMRVLCMLYPL--RAGSSSGCTDPFVIAGRVLTCPSPDFSLFLV 221
DB 173 LRPNHYHTGDPGLVPTLVLSVYNLNPFGGSTS-----AIQVTVETTPSEDPEFVMIR 227
QY 222 PPNVEQTKPFSVNLPLNTLSNSRVPSLIKSMVSRDHCOMVQFQ-----NGRVT 272
DB 228 APS--SKTVDSISAGLLTT-----PVLTVGNDNRWNGQIVGLQVPVGGFSTCNRHWN 279
QY 273 LDGLOQTPTSASOLCKIRGSVFHANGNGYNLTLEL-----DGSPIHAFESP-----APIG 324
DB 280 LNGSTYGWSSPRFADIDHRRGSASY--SGNNATNVLFQFWYANAGS---AIDNPISQVAPDG 335
QY 325 FPDLGECDMHMEASPTTQFNTGDIKQINVKQESAFAPHLGTIOADGLSDSVNTNMIK 384
DB 336 FPDMSFVPFNGPGIPAGWVGFGALWNSN-----SGAPNVTIVQA-----YE 377
QY 385 LGWVSPVSDHGRDVPWIPRYGSLTEAAQ-LAPPIYPPGEGEA 429
DB 378 LGFAT----GAPGNLQP-----TTNTSGAQTVAKSIYAVVTGTA 412

```

Search completed: March 10, 2003, 18:26:16
Job time : 45 secs

